NUMERICAL SOLUTION OF PARABOLIC INTEGRO-DIFFERENTIAL EQUATIONS BY THE DISCONTINUOUS GALERKIN METHOD

STIG LARSSON, VIDAR THOMÉE, AND LARS B. WAHLBIN

Abstract. The numerical solution of a parabolic equation with memory is considered. The equation is first discretized in time by means of the discontinuous Galerkin method with piecewise constant or piecewise linear approximating functions. The analysis presented allows variable time steps which, as will be shown, can then efficiently be selected to match singularities in the solution induced by singularities in the kernel of the memory term or by nonsmooth initial data. The combination with finite element discretization in space is also studied.

1. Introduction

Let H be a separable Hilbert space and assume that A is a linear, selfadjoint, positive definite, not necessarily bounded operator, with compact inverse, defined in $D(A) \subset H$, and that, for $0 \le s < t \le T$, $B(t,s)$ is a linear operator in H with $D(B(t,s)) \supset D(A)$. Consider the initial value problem

$$u_t + Au + \int_0^t B(t,s)u(s)\,ds = f, \quad \text{for } t \in (0,T], \quad \text{with } u(0) = u_0, \tag{1.1}$$

where $f = f(t)$, $u = u(t)$, $u_t = du/dt$. Setting $\|v\|_p = \|A^{p/2}v\| = (A^p v, v)^{1/2}$, where $\|\cdot\|$ is the norm and $(\cdot,\cdot)$ the the inner product in H, we assume throughout the paper that the operator A dominates $B(t,s)$ in the sense that, for some $\alpha \in (0,1]$,

$$|(B(t,s)v,w)| \le C(t-s)^{\alpha-1}\|v\|_p\,\|w\|_q, \quad p = 0,1,2,\ p+q = 2. \tag{1.2}$$

For $0 < \alpha < 1$, (1.2) reflects a weakly singular behavior of $B(t,s)$. When $\alpha = 1$ we shall sometimes assume that an appropriate number of derivatives of $B(t,s)$ exist and are also dominated by A; in this case we refer to $B(t,s)$ as a "smooth kernel".

In the applications that we have in mind, either A is an elliptic second order differential operator in a bounded domain $\Omega \subset \mathbf{R}^d$ with homogeneous Dirichlet boundary conditions, and $B(t,s)$ is a second order differential operator, or else A and $B(t,s)$ are discrete analogs of such operators, arising from a finite element discretization in the spatial variables. Our abstract framework makes it possible to treat these cases simultaneously. In the differential operators case, (1.2) amounts

---

1991 *Mathematics Subject Classification.* 65M60, 65R20, 45L10.

Key words and phrases. Parabolic integro-differential equation, weakly singular kernel, discontinuous Galerkin, variable time step, finite element, error estimate, Gronwall lemma.

The first two authors were partly supported by the Swedish Research Council for Engineering Sciences (TFR). The third author thanks the National Science Foundation, USA, for financial support and also Chalmers University of Technology and Göteborg University for their hospitality during the Spring of 1995.

Published in *Math. Comp.* **67** (1998), 45–71.

to elliptic regularity, plus a bound for the coefficients of $B(t,s)$. The problem considered may, e.g., be thought of as a model problem occurring in the theory of heat conduction in materials with memory, cf. [3]. Equations with weakly singular kernels occur in [7], [9], [10]. For other references, see, e.g., [16].

We shall consider the approximate solution of (1.1) by means of the discontinuous Galerkin method (cf. [4], [5]), which we shall define below in the present context. When A and $B(t,s)$ are differential operators, we shall consider also the discretization in space by finite elements, which will then define a fully discrete method for (1.1) in this case.

For earlier work on discretization in time or space, or both, of equations such as (1.1), see, e.g., [1], [2], [8], [11], [12], [13], [14], [15], [16], [17]. As we shall see in Section 5, a weakly singular kernel in the memory term typically leads to a singularity in the solution (with respect to time), as do nonsmooth initial data. It is hence of interest that the discontinuous Galerkin method, as compared to standard finite difference methods in time, facilitates the analysis of variable time steps and, also, accepts lower regularity of solutions. These advantages of the discontinuous Galerkin method are well known in the context of standard parabolic problems, cf. [4], [5]. However, we point out that our analysis in this paper does not take into account numerical approximation of certain integrals occurring.

To define our time stepping method, let $0 = t_0 < t_1 < \cdots < t_n < \cdots \leq T$ be a partition of the interval $[0,T]$, and define $I_n = (t_{n-1}, t_n)$, $k_n = t_n - t_{n-1}$. Further let $\mathcal{V}_N = \mathcal{V}_N{}^q$, for $t_N \in (0,T]$, denote the set of scalar functions on $[0,t_N]$, which, for $n = 1,\dots,N$, reduce to polynomials of degree less than q on I_n with $q = 1$ or 2. We shall work with functions in $\mathcal{W}_N \equiv \mathcal{V}_N \otimes D(A^{1/2})$; in the differential operator applications these are functions of $(x,t) \in \Omega \times [0,t_N]$, which are either piecewise constant or piecewise linear in time, not necessarily continuous at the nodes of the partition.

Letting $A(v,w)$ and $B(t,s;v,w)$ denote the natural bilinear forms on $D(A^{1/2})$ generated by (Av,w) and $(B(t,s)v,w)$, respectively, we set, for piecewise smooth functions V, W, with $[V]_n = V_n^+ - V_n^-$, $V_n^\pm = \lim_{t\to t_n\pm} V(t)$ denoting jump terms,

$$
\begin{aligned}
G_N(V,W) = \sum_{n=1}^{N} \int_{I_n} \Big(&(V_t(t), W(t)) + A(V(t), W(t)) \\
&+ \int_0^t B(t,s;V(s),W(t))\,ds \Big)\,dt + \sum_{n=1}^{N-1} ([V]_n, W_n^+) + (V_0^+, W_0^+).
\end{aligned} \tag{1.3}
$$

For $B(t,s) \equiv 0$ we recognize the bilinear form used in the analysis of the discontinuous Galerkin method for a parabolic differential equation.

Multiplication in H of (1.1) by X and integration over $(0,t_N)$ show that the exact solution satisfies

$$G_N(u,X) = (u_0, X_0^+) + \int_0^{t_N} (f(t), X(t))\,dt, \quad \forall X \in \mathcal{W}_N.$$

The numerical approximation $U \in \mathcal{W}_N$ is now defined by

$$G_N(U,X) = (u_0, X_0^+) + \int_0^{t_N} (f(t), X(t))\,dt, \quad \forall X \in \mathcal{W}_N. \tag{1.4}$$

We note that this is a time stepping scheme, which determines U successively on I_n for $n = 1, \ldots, N$, when it is known on $[0, t_{n-1})$, from

$$\int_{I_n} \Big((U_t, X) + A(U, X) + \int_{t_{n-1}}^{t} B(\cdot, s; U(s), X)\, ds \Big)\, dt + (U_{n-1}^+, X_{n-1}^+)$$
$$= (U_{n-1}^-, X_{n-1}^+) + \int_{I_n} (f, X)\, dt - \int_{I_n} \int_0^{t_{n-1}} B(\cdot, s; U(s), X)\, ds\, dt, \quad \forall X \in \mathcal{W}_N,$$

where $U_0^- = u_0$. The uniqueness of U follows by Gronwall's lemma provided that $k = \max_n k_n$ is small enough (which we shall assume in the sequel without specific mention), cf. the stability estimate (3.2) in Theorem 3.1 below. The existence in the case $B(t, s) \equiv 0$ follows from the uniqueness since, using the eigenspaces of A, (1.1) can then be reduced to finite dimensional problems. For $B(t, s) \not\equiv 0$ and k small, the problem may be thought of as a small perturbation of the problem with $B(t, s) \equiv 0$, which may be solved by the contraction mapping theorem.

We note that $U - u$ satisfies the "orthogonality" condition

$$G_N(U - u, X) = 0, \quad \forall X \in \mathcal{W}_N. \tag{1.5}$$

Our first error estimate is now as follows. Here and below we set

$$|g|_{I_n} = \sup_{I_n} \|g(t)\| \quad \text{and} \quad |g|_{p, I_n} = \sup_{I_n} \|g(t)\|_p.$$

We shall often also use the analogous notation $|g|_{J_N}$ and $|g|_{p, J_N}$, where $J_N = (0, t_N)$, and write D_t for d/dt.

Theorem 1.1. *Let U and u be the solutions of (1.4) and (1.1). Then there exists a constant $C = C(T)$ such that, for $t_N \in [0, T]$,*

$$|U - u|_{I_N} \le C k_N^q |D_t^q u|_{I_N} + C \sum_{n=1}^{N} k_n^{q+1} |D_t^q u|_{2, I_n}.$$

In particular, this error bound is of order $O(k^q)$ for a smooth solution u.

Although the error bound derived in Theorem 1.1 does point at the interplay between the regularity of u and good choices of the time steps, it has a form which makes an explicit choice difficult. For this reason we now present an estimate where the l_1-norm in time has been replaced by a maximum norm. Here and below, we denote $L_N = (1 + \log(t_N/k_N))^{1/2}$, which is of moderate size compared to $1/k_N$, and we shall always assume the mesh ratio condition $k_n/k_{n+1} \le \omega$, for $n \ge 1$.

Theorem 1.2. *Let U and u be the solutions of (1.4) and (1.1). Then there exists a constant $C = C(T)$ such that, for $t_N \in [0, T]$,*

$$\|U_N^- - u(t_N)\| \le C L_N \max_{1 \le n \le N} \big(k_n^q |D_t^q u|_{I_n} \big).$$

When $q = 2$ it is clear that if we interpolate linearly in I_n between U_{n-1}^- and U_n^- we will obtain an approximation for $u(t)$ on all of I_n with the same error bound as in Theorem 1.2; for a smooth solution this thus shows a global second order error bound. We shall next give a result for the piecewise linear case, which shows that then (under appropriate smoothness assumptions) the error in the nodal value U_N^- is of superconvergent third order in k.

Theorem 1.3. *Let $q = 2$ and let U and u be the solutions of (1.4) and (1.1). If either $B(t,s)$ is smooth and such that also $B_t(t,s)$ is dominated by A, or if $B(t,s)$ is of convolution type, i.e., $B(t,s) = B(t-s)$, then there exists a constant $C = C(T)$ such that, for $t_N \in [0,T]$,*

$$\|U_N^- - u(t_N)\| \le CL_N \max_{1\le n\le N} \big(k_n^3 |D_t^2 u|_{2,I_n}\big).$$

We remark that in the quantities such as $k_n^q|D_t^q u|_{I_n}$ appearing in Theorems 1.1–1.3, the number q may be replaced by any m with $0 \le m \le q$; with the proper interpretation m may also be fractional. Note also that, in all of our results, the variable t_N is allowed to vary in the interval $[0,T]$. Due to the use of Gronwall's lemma, the quantities $C(T)$ in our estimates grow rapidly as T increases, cf. [13].

The proofs of our error estimates will be carried out by energy arguments; for Theorems 1.2 and 1.3 they will depend on the following stability result for the dual problem to (1.4) with $f = 0$.

Theorem 1.4. *Let $t_N \in [0,T]$ and let $Z \in \mathcal{W}_N$ be given by*

$$G_N(X,Z) = (X_N^-, \varphi), \quad \forall X \in \mathcal{W}_N, \tag{1.6}$$

where $\varphi \in H$. Then there exists a constant $C = C(T)$ such that

$$|Z|_{J_N} \le C\|\varphi\|,$$

and, with $[Z]_N = \varphi - Z_N^-$,

$$\sum_{n=1}^N \Big(\int_{I_n} \big(\|Z_t\| + \|Z\|_2\big)\, dt + \|[Z]_n\| \Big) \le CL_N\|\varphi\|.$$

Note that (1.6) is a discrete version of the backward evolution problem

$$-z_t + Az + \int_t^{t_N} B^*(s,t) z(s)\, ds = 0, \quad \text{for } t \in (0,t_N), \quad \text{with } z(t_N) = \varphi, \tag{1.7}$$

where $B^*(s,t)$ is the adjoint of $B(s,t)$. The proof of Theorem 1.4 is carried out in Theorem 3.1 for a related forward equation of the form (1.4) with $u_0 = \varphi$, $f = 0$. It is known that Theorem 1.4 holds in the case that $B(t,s) \equiv 0$, i.e., when no memory term is present (see [4], Lemma 6.1). In our case we therefore write the solution in the form $U = V + W$, where V is the solution for $B(t,s) \equiv 0$. It then remains to show the estimates of Theorem 1.4 for W, which satisfies

$$\sum_{n=1}^N \int_{I_n} \Big((W_t, X) + A(W,X) \Big)\, dt + \sum_{n=1}^{N-1} ([W]_n, X_n^+) + (W_0^+, X_0^+)$$
$$= -\int_0^{t_N} \int_0^t B(\cdot, s; U(s), X)\, ds\, dt, \quad \forall X \in \mathcal{W}_N.$$

The main part of the proof now consists in showing

$$\int_0^{t_N} \|W\|_2\, dt \le C(T)\|\varphi\|. \tag{1.8}$$

When $B(t,s)$ is smooth, and also for singular kernels with $\alpha > \frac12$, this follows at once from showing that $\|W\|_2$ is square integrable in time, which is proved by a simple energy argument. For $\alpha \le \frac12$ more technical and somewhat lengthy considerations show that $t^\gamma\|W(t)\|_2$ is square integrable with $\gamma < \frac12$, which again implies (1.8).

We next give some simple examples of how Theorems 1.1–1.3 apply, particularly with reference to the regularity of the solution and the choice of the time steps. We start with $B(t,s)$ smooth. Then Theorem 1.1 shows a $O(k^q)$ error bound provided $\|D_t^q u(t)\|_2$ is bounded, whereas Theorem 1.2 gives $O(L_N k^q)$, assuming only that $\|D_t^q u(t)\|$ is bounded. Conditions on u_0 and f for such properties to hold are given in [17], Theorem 2.3. In particular, for the homogeneous equation ($f = 0$), when $q = 1$, the requirements are $u_0 \in D(A^2)$ and $u_0 \in D(A)$, respectively, and when $q = 2$, additionally that $(A^2 - B(0,0))u_0 \in D(A)$ and $u_0 \in D(A^2)$, respectively. In these cases there are, a priori, no pressing reasons to use variable time steps.

Consider then the limited regularity case when $u_0 \in D(A)$ only, still with $B(t,s)$ smooth and $f = 0$. It may then be shown, using the techniques in [14], [15], that

$$\|u_t(t)\| + \|u(t)\|_2 + t\|u_{tt}(t)\| + t\|u_t(t)\|_2 + t^2\|u_{tt}(t)\|_2 \le C_0, \quad t \in (0,T], \tag{1.9}$$

so that the bounds for the higher derivatives are singular at $t = 0$. Then, in the piecewise constant case, Theorem 1.1 together with (1.9) shows

$$|U - u|_{I_N} \le CC_0\Big(k_N + k_1 + \sum_{n=2}^{N} k_n^2/t_{n-1}\Big) \le CC_0 L_N^2 k \tag{1.10}$$

(we always assume that k_n is chosen so that $k_n \le t_{n-1}$, for $n \ge 2$), while Theorem 1.2 similarly gives the slightly smaller bound

$$\|U_N^- - u(t_N)\| \le CC_0 L_N \max_{1\le n\le N} k_n = CC_0 L_N k. \tag{1.11}$$

For piecewise linear functions Theorem 1.2 gives

$$\|U_N^- - u(t_N)\| \le CC_0 L_N\Big(k_1 + \max_{2\le n\le N}(k_n^2/t_{n-1})\Big), \tag{1.12}$$

while Theorem 1.3 shows the higher order estimate

$$\|U_N^- - u(t_N)\| \le CC_0 L_N\Big(k_1 + \max_{2\le n\le N}(k_n^3/t_{n-1}^2)\Big). \tag{1.13}$$

Let us now turn to the case of a weakly singular kernel, which, as already noted, accounts for much of our technical analysis in this paper. We start with the special case of a kernel of the form $B(t,s) = (t-s)^{\alpha-1}B$ and, assume first that $u_0 \in D(A)$ only (and that $f = 0$). As we shall see below in Section 5, the estimates of (1.9) still hold, except the one for $\|u_{tt}\|_2$, and hence (1.10), (1.11), and (1.12) remain valid. As for Theorem 1.3, if $B = A$ and if $u_0 \in D(A^{2+\alpha})$, then $\|u_{tt}(t)\|_2 \le Ct^{\alpha-1}$, and an analysis similar to the above is easily furnished.

We shall see below that, in fact, the second derivatives deteriorate near $t = 0$, no matter how smooth the initial data u_0 are. If $u_0 \in D(A^{1+\alpha})$, then $\|u_{tt}(t)\| \le Ct^{\alpha-1}$ (and $u_t \in C^\alpha(\bar J_N, L_2)$) so that, for piecewise linear functions, we have by Theorem 1.2 (with a fractional power of k_1 on I_1),

$$\|U_N^- - u(t_N)\| \le CC_0 L_N\Big(k_1^{1+\alpha} + \max_{2\le n\le N}(k_n^2/t_{n-1}^{1-\alpha})\Big). \tag{1.14}$$

Estimates like (1.13) and (1.14) give guidelines on how to choose suitable mesh refinements. For instance, with $T = t_M = 1$, choosing the time levels by $t_n = (n/M)^\gamma$, $n = 0, 1, \ldots, M$, we find $k_n \approx \gamma M^{-1} t_{n-1}^{(\gamma-1)/\gamma}$. Then taking $\gamma = 3$ we find from (1.13) that $\|U_N^- - u(t_N)\| \le CC_0 L_N M^{-3}$ for $N \le M$, while $\gamma = 2/(1+\alpha)$, in (1.14) gives $\|U_N^- - u(t_N)\| \le CC_0 L_N M^{-2}$.

We next consider the case when A and $B(t,s)$ are differential operators of the form described above, and take $H = L_2(\Omega)$. We shall then study a fully discrete method combining time stepping by the discrete Galerkin method introduced above with the use of finite elements for the approximation in the spatial variables. We assume that we are given a family of finite dimensional spaces $S_h \subset H_0^1(\Omega)$ such that there is a positive integer r, the order of accuracy of $\{S_h\}$, such that, with $H^p = H^p(\Omega)$ the standard L_2-based Sobolev spaces,

$$\inf_{\chi \in S_h} \big(\|v-\chi\| + h\|\nabla(v-\chi)\|\big) \le Ch^p\|v\|_{H^p}, \quad 1 \le p \le r,\ v \in H^p \cap H_0^1. \tag{1.15}$$

The approximation U of the solution of (1.1) is now sought in $\mathcal{W}_{N,h} \equiv \mathcal{V}_N \otimes S_h$ from

$$G_N(U,X) = (u_{0h}, X_0^+) + \int_0^{t_N} (f,X)\,dt, \quad \forall X \in \mathcal{W}_{N,h}, \tag{1.16}$$

with $G_N(\cdot,\cdot)$ defined as above, and where $u_{0h} \in S_h$ is a given approximation to u_0. The error equation is in this case, cf. (1.5),

$$G_N(U-u,X) = (u_{0h} - u_0, X_0^+), \quad \forall X \in \mathcal{W}_{N,h}. \tag{1.17}$$

In particular, the right hand side vanishes if $u_{0h} = P_h u_0$, where P_h denotes the L_2-projection onto S_h.

As examples of results in this case we shall present analogs of Theorems 1.2 and 1.3 above. In both cases our proofs will involve domination for discrete analogs of A and $B(t,s)$. We therefore make the following assumption, which will serve to transfer domination, i.e., (1.2) from the continuous to the discrete case: Either

$$B(t,s) = b(t,s)A + A_0(t,s), \tag{1.18}$$

where $b(t,s)$ is a scalar function and $A_0(t,s)$ is a differential operator of at most first order, or the finite element spaces have an inverse property

$$\|\chi\|_1 \le Ch^{-1}\|\chi\|, \quad \forall \chi \in S_h. \tag{1.19}$$

Theorem 1.5. *Let U and u be the solutions of (1.16) and (1.1), respectively, and assume that (1.15) and (1.18) or (1.19) holds. Then there exists a constant $C = C(T)$ such that, for $t_N \in [0,T]$,*

$$\|U_N^- - u(t_N)\| \le CL_N\Big(\|u_{0h} - u_0\| + h^p \sup_{0\le t\le t_N} \|u(t)\|_{H^p} + \max_{1\le n\le N} \big(k_n^q |D_t^q u|_{I_n}\big)\Big).$$

Theorem 1.6. *Let $q = 2$, let U and u be the solutions of (1.16) and (1.1), and assume that (1.15) and (1.18) or (1.19) holds. If $B(t,s)$ is smooth and such that also $B_t(t,s)$ and $B_{tt}(t,s)$ are dominated by A, then there exists a constant $C = C(T)$ such that, for $t_N \in [0,T]$,*

$$\begin{aligned}\|U_N^- - u(t_N)\| \le CL_N\Bigg(&\|u_{0h} - u_0\| + h^p \sup_{0\le t\le t_N} \|u(t)\|_{H^p} \\ &+ \max_{1\le n\le N} \Big\{k_n^3\Big(\sum_{l=0}^{2} |D_t^l u|_{2,I_n} + \int_0^{t_n} \|u\|_2\,ds\Big)\Big\}\Bigg).\end{aligned}$$

The plan of the paper is as follows: In Section 2 we show the error estimates of Theorems 1.1–1.3, assuming the stability result of Theorem 1.4, which is proved in Section 3. The completely discrete case with finite elements in the spatial variables is analyzed in Section 4 and some examples of the use of the error estimates are given. Finally, some examples of regularity results of the type used above are shown in Section 5. Some technical lemmas are collected in Section 6.

We finally remark that on several occasions we claim existence of solutions of our problems in specified functions spaces, as a result of the corresponding a priori estimates. Such results can be justified using the Faedo-Galerkin method based on eigenvector expansions associated with the operator A, cf., e.g., [2], Theorem 1, for an example in the context of integro-differential equations.

2. Proofs of the error estimates

In this section we prove the error estimates of Theorems 1.1, 1.2, and 1.3. The proofs of the latter two results are based on the stability estimates of Theorem 1.4. In all of the proofs we use the linear interpolation operator Π, which maps smooth functions of t onto $\mathcal{V}_N$, and which is defined, for $n = 1, \dots, N$, by

$$(\Pi g)_n^- = g(t_n), \quad \text{if } q = 1, 2, \quad \text{and} \int_{I_n} (\Pi g)(t)\, dt = \int_{I_n} g(t)\, dt, \quad \text{if } q = 2.$$

Then Π approximates the identity operator to order q, i.e.,

$$|\Pi g - g|_{I_n} \le C k_n^m |D_t^m g|_{I_n}, \quad \text{for } 0 \le m \le q. \tag{2.1}$$

Note in particular the case $m = 0$, which means that Π is stable with respect to $|\cdot|_{I_n}$. Writing

$$U - u = (U - \Pi u) + (\Pi u - u) \equiv \theta + \eta,$$

we thus have access to bounds for η. For the other term $\theta \in \mathcal{W}_N$ we note that the "orthogonality" relation (1.5) yields

$$G_N(\theta, X) = -G_N(\eta, X), \quad \forall X \in \mathcal{W}_N,$$

where, by integration by parts in definition (1.3),

$$\begin{aligned} G_N(V,W) = \sum_{n=1}^{N} \int_{I_n} \Big(-(V, W_t) + A(V,W) + \int_0^t B(\cdot, s; V(s), W)\, ds \Big)\, dt \\ - \sum_{n=1}^{N-1} (V_n^-, [W]_n) + (V_N^-, W_N^-). \end{aligned} \tag{2.2}$$

Using also the defining properties of Π, we conclude that $\theta \in \mathcal{W}_N$ satisfies the equation

$$G_N(\theta, X) = -\int_0^{t_N} \Big(A(\eta, X) + \int_0^t B(\cdot, s; \eta(s), X)\, ds \Big)\, dt, \quad \forall X \in \mathcal{W}_N. \tag{2.3}$$

Proof of Theorem 1.1. By (2.1), η is bounded as desired. To estimate θ we choose $X = \theta$ in (2.3), and after a simple calculation we obtain

$$\begin{aligned} \tfrac12 \|\theta_N^-\|^2 + \tfrac12 \|\theta_0^+\|^2 + \tfrac12 \sum_{n=1}^{N-1} \|[\theta]_n\|^2 + \int_0^{t_N} \|\theta\|_1^2\, dt = -\int_0^{t_N} A(\eta, \theta)\, dt \\ - \int_0^{t_N} \int_0^t B(\cdot, s; \eta(s), \theta)\, ds\, dt - \int_0^{t_N} \int_0^t B(\cdot, s; \theta(s), \theta)\, ds\, dt \\ \equiv R_1 + R_2 + R_3. \end{aligned} \tag{2.4}$$

Here, $R_1 \leq \int_0^{t_N} \|\eta\|_2\, dt\, |\theta|_{J_N}$, and, by assumption (1.2) and a change of order of integration,

$$R_2 \leq C \int_0^{t_N} \int_0^t (t-s)^{\alpha-1} \|\eta(s)\|_2\, ds\, \|\theta(t)\|\, dt \leq C \int_0^{t_N} \|\eta\|_2\, dt\, |\theta|_{J_N}.$$

Further, using also Schwarz' inequality and in the last step the technical estimate of Lemma 6.3 (with $\delta = 0$), combined with the arithmetic-geometric mean inequality, we have

$$\begin{aligned} R_3 &\leq C \int_0^{t_N} \int_0^t (t-s)^{\alpha-1} \|\theta(s)\|_1\, ds\, \|\theta(t)\|_1\, dt \\ &\leq C \Big(\int_0^{t_N} \Big(\int_0^t (t-s)^{\alpha-1} \|\theta(s)\|_1\, ds \Big)^2 dt \Big)^{1/2} \Big(\int_0^{t_N} \|\theta\|_1^2\, dt \Big)^{1/2} \\ &\leq C \int_0^{t_N} (t_N - t)^{\alpha-1} \int_0^t \|\theta(s)\|_1^2\, ds\, dt + \tfrac{1}{2} \int_0^{t_N} \|\theta\|_1^2\, dt. \end{aligned}$$

Clearly, since $q = 1, 2$, we have (recall that $|v|_{J_N} = \sup_{t \in (0,t_N)} \|v(t)\|$)

$$|\theta|_{J_N} \leq \max_{1 \leq n \leq N} (\|\theta_n^-\| + \|\theta_{n-1}^+\|) \leq 2 \max_{1 \leq n \leq N} \|\theta_n^-\| + \max_{1 \leq n \leq N-1} \|[\theta]_n\| + \|\theta_0^+\|,$$

and hence

$$|\theta|_{J_N}^2 \leq C \max_{1 \leq n \leq N} \Big(\|\theta_n^-\|^2 + \sum_{n=1}^{N-1} \|[\theta]_n\|^2 + \|\theta_0^+\|^2 \Big).$$

We therefore conclude from (2.4) and the above that

$$|\theta|_{J_N}^2 + \int_0^{t_N} \|\theta\|_1^2\, dt \leq C \Big(\int_0^{t_N} \|\eta\|_2\, dt \Big)^2 + C \int_0^{t_N} (t_N - t)^{\alpha-1} \int_0^t \|\theta\|_1^2\, ds\, dt.$$

Denoting the left side by ϕ_N and the first term on the right by a_N, we have, since $t \leq t_n$ for $t \in I_n$,

$$\begin{aligned} \phi_N &\leq a_N + C \sum_{n=1}^N \int_{I_n} (t_N - t)^{\alpha-1}\, dt \int_0^{t_n} \|\theta\|_1^2\, ds \\ &\leq a_N + C \sum_{n=1}^N \int_{I_n} (t_N - t)^{\alpha-1}\, dt\ \phi_n. \end{aligned}$$

Using a variant of Gronwall's lemma (Lemma 6.4 below) and (2.1), we find

$$|\theta|_{J_N} \leq C \int_0^{t_N} \|\eta\|_2\, dt \leq C \sum_{n=1}^N k_n |(\Pi - I)u|_{2,I_n} \leq C \sum_{n=1}^N k_n^{q+1} |D_t^q u|_{2,I_n}.$$

Since $U - u = \theta + \eta$, this completes the proof. ☐

Proof of Theorem 1.2. We shall estimate $U_N^- - u(t_N)$ by duality. Let Z be the solution of (1.6) with $\|\varphi\| = 1$. Since $\eta_N^- = 0$ we have

$$(U_N^- - u(t_N), \varphi) = (\theta_N^-, \varphi) = G_N(\theta, Z),$$

so that, in view of (2.3),

$$(U_N^- - u(t_N), \varphi) = -\int_0^{t_N} \Big(A(\eta, Z) + \int_0^t B(\cdot, s; \eta(s), Z)\, ds \Big)\, dt. \tag{2.5}$$

Using assumption (1.2) and Theorem 1.4, we obtain

$$|(U_N^- - u(t_N), \varphi)| \le C|\eta|_{J_N} \int_0^{t_N} \|Z\|_2 \, dt \le CL_N |\eta|_{J_N},$$

which proves the theorem. □

Proof of Theorem 1.3. Again we use the error representation (2.5) and Theorem 1.4. Switching the order of integration in the memory term in (2.5) yields

$$\begin{aligned}(U_N^- - u(t_N), \varphi) &= -\int_0^{t_N} \Big(A(\eta, Z) + \int_t^{t_N} B(s,t;\eta(t), Z(s))\, ds \Big)\, dt \\ &= -\int_0^{t_N} \Big(A\eta(t), Z(t) + \int_t^{t_N} A^{-1}B^*(s,t)Z(s)\, ds \Big)\, dt = -\sum_{n=1}^{N} \int_{I_n} (A\eta, K)\, dt,\end{aligned}$$

where

$$K(t) = Z(t) + \int_t^{t_N} A^{-1}B^*(s,t)Z(s)\, ds.$$

Since η is orthogonal to constants (recall that $q = 2$), we have

$$\Big| \int_{I_n} (A\eta, K)\, dt \Big| = \Big| \int_{I_n} \Big(A\eta(t), K_{n-1}^+ + \int_{t_{n-1}}^{t} K_t\, ds \Big)\, dt \Big| \le k_n |\eta|_{2,I_n} \int_{I_n} \|K_t\|\, ds.$$

Hence

$$|(U_N^- - u(t_N), \varphi)| \le \max_{1 \le n \le N} \big(k_n |\eta|_{2,I_n} \big) \sum_{n=1}^{N} \int_{I_n} \|K_t\|\, dt.$$

Since $\eta = \Pi u - u$ may be estimated by (2.1), it remains to prove

$$\sum_{n=1}^{N} \int_{I_n} \|K_t\|\, dt \le CL_N. \tag{2.6}$$

If $B(t,s)$ is smooth, then, for $t \in I_n$,

$$K_t(t) = Z_t(t) - A^{-1}B^*(t,t)Z(t) + \int_t^{t_N} A^{-1}B_t^*(s,t)Z(s)\, ds.$$

We note that (1.2) gives $\|A^{-1}B^*(t,t)\| \le C$ and similarly, under our present assumptions, $\|A^{-1}B_t^*(s,t)\| \le C$, and hence (2.6) easily follows from Theorem 1.4.

If the kernel is of convolution type with $B(t,s) = B(t-s)$, then

$$K(t) = Z(t) + A^{-1} \int_t^{t_N} B^*(s-t)Z(s)\, ds = Z(t) + A^{-1} \int_0^{t_N - t} B^*(s)Z(s+t)\, ds.$$

It follows by a direct calculation that K is differentiable for $t \ne t_n$, $n = 0, \dots, N$. In fact, for $t \in I_n$, we have

$$K_t(t) = Z_t(t) + A^{-1} \Big(\int_t^{t_N} B^*(s-t)Z_t(s)\, ds + \sum_{l=n}^{N-1} B^*(t_l - t)[Z]_l - b(t_N - t)Z_N^- \Big),$$

where Z_t denotes the piecewise constant function obtained by differentiation of Z, and the sum is empty if $n = N$. Noting that (1.2) implies $\|A^{-1}B^*(t)\| \le Ct^{\alpha-1}$, we get

$$\sum_{n=1}^{N}\int_{I_n}\|K_t\|\,dt \le \int_0^{t_N}\|Z_t\|\,dt + C\int_0^{t_N}\int_t^{t_N}(s-t)^{\alpha-1}\|Z_t(s)\|\,ds\,dt$$
$$+\,C\sum_{n=1}^{N-1}\int_{I_n}\sum_{l=n}^{N-1}(t_l-t)^{\alpha-1}\,dt\,\|[Z]_l\| + C\int_0^{t_N}(t_N-t)^{\alpha-1}\,dt\,\|Z_N^-\|,$$

and after changing the order of integration and summation,

$$\sum_{n=1}^{N}\int_{I_n}\|K_t\|\,dt \le C\int_0^{t_N}\|Z_t\|\,dt + C\sum_{l=1}^{N-1}\|[Z]_l\|\sum_{n=1}^{l}\int_{I_n}(t_l-t)^{\alpha-1}\,dt + C\|Z_N^-\|$$
$$\le C\Big(\int_0^{t_N}\|Z_t\|\,dt + \sum_{n=1}^{N-1}\|[Z]_n\| + \|Z_N^-\|\Big),$$

which proves (2.6) and thus completes the proof. □

3. Stability

The object of this section is to prove our main stability result, Theorem 1.4, for the backward evolution problem (1.6). It is convenient to carry out the proof for a related forward problem obtained by a change of variable $t \to t_N - t$. More precisely, setting $\tilde z(t) = z(t_N - t)$ in (1.7) yields

$$\tilde z_t + A\tilde z + \int_0^t \tilde B(t,s)\tilde z(s)\,ds = 0,\ t\in(0,t_N);\quad \tilde z(0)=\varphi,$$

where $\tilde B(t,s) = B^*(t_N - s, t_N - t)$. A similar consideration applies to (1.6), which is thus equivalent to a forward problem of the form (1.4) with $f = 0$, $u_0 = \varphi$, with a different kernel, namely $\tilde B(t,s)$, and a reversed mesh. By noting that $\tilde B(t,s)$ satisfies (1.2), and inverting the mesh ratio condition, we see that Theorem 1.4 follows from the following theorem, where $\tilde L_N = (1+\log(t_N/k_1))^{1/2}$.

Theorem 3.1. *Let $t_N \in [0,T]$ and let $U \in \mathcal{W}_N$ be defined by*

$$G_N(U,X) = (\varphi, X_0^+),\quad \forall X\in\mathcal{W}_N, \tag{3.1}$$

where $\varphi \in H$. Then there exists a constant $C = C(T)$ such that

$$|U|_{J_N} \le C\|\varphi\|, \tag{3.2}$$

and, if $k_{n+1}/k_n \le \omega$ for $n \ge 1$, a constant $C = C(T,\omega)$ such that, with $[U]_0 = U_0^+ - \varphi$,

$$\sum_{n=1}^{N}\Big(\int_{I_n}\big(\|U_t\| + \|U\|_2\big)\,dt + \|[U]_{n-1}\|\Big) \le C\tilde L_N\|\varphi\|. \tag{3.3}$$

We begin by proving (3.2).

Proof of (3.2). With $X = U$ in (3.1) we have

$$\tfrac12\|U_N^-\|^2 + \tfrac12\|U_0^+\|^2 + \tfrac12\sum_{n=1}^{N-1}\|[U]_n\|^2 + \int_0^{t_N}\|U\|_1^2\,dt$$
$$= (\varphi, U_0^+) - \int_0^{t_N}\int_0^t B(t,s;U(s),U(t))\,ds\,dt.$$

Hence, since $(\varphi, U_0^+) = \frac{1}{2}\big(\|\varphi\|^2 + \|U_0^+\|^2 - \|[U]_0\|^2\big)$, and in view of (1.2), we get

$$\begin{aligned}
\|U_N^-\|^2 &+ \sum_{n=0}^{N-1} \|[U]_n\|^2 + 2\int_0^{t_N} \|U\|_1^2\,dt \\
&\le \|\varphi\|^2 + C\int_0^{t_N} \|U(t)\|_1 \int_0^t (t-s)^{\alpha-1}\|U(s)\|_1\,ds\,dt \\
&\le \|\varphi\|^2 + \int_0^{t_N} \|U\|_1^2\,dt + C\int_0^{t_N} \Big(\int_0^t (t-s)^{\alpha-1}\|U(s)\|_1\,ds\Big)^2\,dt.
\end{aligned}$$

By application of Lemma 6.3 to the last term we obtain

$$\begin{aligned}
\|U_N^-\|^2 &+ \sum_{n=0}^{N-1} \|[U]_n\|^2 + \int_0^{t_N} \|U\|_1^2\,dt \\
&\le \|\varphi\|^2 + CT^\alpha \int_0^{t_N} (t_N - t)^{\alpha-1} \int_0^t \|U(s)\|_1^2\,ds\,dt,
\end{aligned}$$

and Gronwall's lemma (Lemma 6.4) gives (as in the proof of Theorem 1.1)

$$\|U_N^-\|^2 + \sum_{n=0}^{N-1} \|[U]_n\|^2 + \int_0^{t_N} \|U\|_1^2\,dt \le C(T)\|\varphi\|^2. \tag{3.4}$$

Hence, we have estimated the nodal values U_N^- in the desired way, and the proof will be complete once we have proved

$$\int_{I_n} \|U_t\|\,dt \le C(T)\|\varphi\|, \quad \text{for } 1 \le n \le N. \tag{3.5}$$

In order to show (3.5) we take $X(t) = (t-t_{n-1})U_t(t)$ for $t \in I_n$, $X(t) = 0$ otherwise, in (3.1) to obtain

$$\begin{aligned}
\int_{I_n} (t-t_{n-1})\|U_t(t)\|^2\,dt &= -\int_{I_n} (t-t_{n-1})A(U(t), U_t(t))\,dt \\
&\quad - \int_{I_n} (t-t_{n-1}) \int_0^t B(t,s;U(s),U_t(t))\,ds\,dt \equiv R_1 + R_2.
\end{aligned} \tag{3.6}$$

Here, for R_1 we have, by means of an integration by parts and (3.4),

$$R_1 = -\tfrac{1}{2}\int_{I_n} (t-t_{n-1})\frac{d}{dt}\|U(t)\|_1^2\,dt = -\tfrac{1}{2}k_n\|U_n^-\|_1^2 + \tfrac{1}{2}\int_{I_n} \|U\|_1^2\,dt \le C(T)\|\varphi\|^2.$$

For R_2 we use (1.2), Schwarz' inequality, an inverse inequality, and Young's inequality (Lemma 6.1), to obtain the estimate

$$\begin{aligned}
R_2 &\le C\int_{I_n} (t-t_{n-1})\|U_t(t)\|_1 \int_0^t (t-s)^{\alpha-1}\|U(s)\|_1\,ds\,dt \\
&\le Ck_n^2 \int_{I_n} \|U_t\|_1^2\,dt + C\int_{I_n} \Big(\int_0^t (t-s)^{\alpha-1}\|U(s)\|_1\,ds\Big)^2\,dt \\
&\le C\int_{I_n} \|U\|_1^2\,dt + CT^{2\alpha}\int_0^{t_n} \|U\|_1^2\,dt \le C(T)\int_0^{t_n} \|U\|_1^2\,dt \le C(T)\|\varphi\|^2.
\end{aligned}$$

Since U_t is constant on I_n, we have

$$\Big(\int_{I_n} \|U_t\|\,dt\Big)^2 = 2\int_{I_n} (t-t_{n-1})\|U_t(t)\|^2\,dt, \tag{3.7}$$

and hence the above estimates prove (3.5). ☐

The remaining result (3.3) will be proved by splitting U into two parts $U = V+W$, where $V \in \mathcal{W}_N$ is defined by the purely parabolic discrete problem obtained by setting $B = 0$, i.e., $\forall X \in \mathcal{W}_N$,

$$\sum_{n=1}^{N} \int_{I_n} \Big((V_t, X) + A(V, X)\Big)\, dt + \sum_{n=1}^{N-1} ([V]_n, X_n^+) + (V_0^+, X_0^+) = (\varphi, X_0^+). \tag{3.8}$$

For the "parabolic part" V of U we have the following result. In particular, (3.3) holds with U replaced by V. Recall that $\tilde{L}_N = (1+\log(t_N/k_1))^{1/2}$ and $k_{n+1}/k_n \le \omega$ for $n \ge 1$.

Lemma 3.1. *With $[V]_0 = V_0^+ - \varphi$ we have, for the solution of (3.8),*

$$\|V_N^-\|^2 + 2\int_0^{t_N} \|V\|_1^2\, dt + \sum_{n=1}^{N} \|[V]_{n-1}\|^2 = \|\varphi\|^2,$$

$$\sum_{n=1}^{N} t_n \Big(\int_{I_n} \big(\|V_t\|^2 + \|V\|_2^2\big)\, dt + k_n^{-1} \|[V]_{n-1}\|^2 \Big) \le C(\omega) \|\varphi\|^2,$$

and

$$\sum_{n=1}^{N} \Big(\int_{I_n} \big(\|V_t\| + \|V\|_2\big)\, dt + \|[V]_{n-1}\| \Big) \le C(\omega) \tilde{L}_N \|\varphi\|.$$

We refer to Lemma 6.1 of [4] for the proof of this lemma. We now turn to the estimates for W which is the solution of

$$\begin{aligned} &\sum_{n=1}^{N} \int_{I_n} \Big((W_t, X) + A(W, X)\Big)\, dt + \sum_{n=1}^{N-1} ([W]_n, X_n^+) + (W_0^+, X_0^+) \\ &\qquad = -\int_0^{t_N} \int_0^t B(t, s; U(s), X(t))\, ds\, dt, \quad \forall X \in \mathcal{W}_N. \end{aligned} \tag{3.9}$$

Our aim is to show first that

$$\int_0^{t_N} \|W\|_2\, dt \le C(T, \omega) \tilde{L}_N \|\varphi\|. \tag{3.10}$$

In the case of a smooth kernel ($\alpha = 1$ in (1.2)), and also in the case of a singular kernel with $\alpha > \frac{1}{2}$, this will follow at once from

$$\int_0^{t_N} \|W\|_2^2\, dt \le C(T) \tilde{L}_N^2 \|\varphi\|^2, \tag{3.11}$$

which we shall derive in a simple way in Lemmas 3.2 and 3.3. Since, as follows from the discussion in Section 5 below, $\|w\|_2 = O(t^{\alpha-1})$ where w is the continuous in time analog of W, we do not expect (3.11) to hold for $\alpha \le \frac{1}{2}$, and we shall therefore show instead essentially that $t^\gamma AW(t) \in L_2(J_N, H)$, with $\gamma < \frac{1}{2}$, which implies (3.10). We begin with the case of a smooth kernel.

Lemma 3.2. *Let W be the solution of (3.9). If $B(t, s)$ is a smooth kernel, then*

$$\int_0^{t_N} \|W\|_2\, dt \le C(T) \tilde{L}_N \|\varphi\|.$$

Proof. We shall show (3.11) which clearly implies the desired result. With $X(t) = AW(t)$ in (3.9) we have, by a straightforward calculation,

$$\begin{aligned}\tfrac{1}{2}\|W_N^-\|_1^2 + \tfrac{1}{2}\|W_0^+\|_1^2 + \tfrac{1}{2}\sum_{n=1}^{N-1}\|[W]_n\|_1^2 + \int_0^{t_N}\|W\|_2^2\,dt \\ = -\int_0^{t_N}\int_0^t B(\cdot,s;U(s),AW)\,ds\,dt.\end{aligned}$$

Using the bound for B in (1.2) with $\alpha = 1$, and Schwarz' inequality, the term on the right may be bounded by

$$C\int_0^{t_N}\|W(t)\|_2\int_0^t\|U(s)\|_2\,ds\,dt \le \tfrac{1}{2}\int_0^{t_N}\|W\|_2^2\,dt + C\int_0^{t_N}\Big(\int_0^t\|U\|_2\,ds\Big)^2\,dt.$$

Since $U = V + W$ we hence have

$$\int_0^{t_N}\|W\|_2^2\,dt \le CT\Big(\int_0^{t_N}\|V\|_2\,dt\Big)^2 + CT\int_0^{t_N}\int_0^t\|W\|_2^2\,ds\,dt. \tag{3.12}$$

In view of Lemma 3.1 the first term on the right is bounded by $C(T)\tilde{L}_N^2\|\varphi\|^2$, and by an obvious estimate for the second term, we now have

$$\int_0^{t_N}\|W\|_2^2\,dt \le C(T)\tilde{L}_N^2\|\varphi\|^2 + C(T)\sum_{n=1}^{N}\Big(k_n\int_0^{t_n}\|W\|_2^2\,dt\Big).$$

The desired result now follows by the standard discrete Gronwall lemma. □

Lemma 3.3. *The conclusion of Lemma 3.2 remains valid if $B(t,s)$ is weakly singular with $\alpha > \frac{1}{2}$.*

Proof. In this case we obtain instead of (3.12)

$$\begin{aligned}\int_0^{t_N}\|W\|_2^2\,dt \le C\int_0^{t_N}\Big(\int_0^t (t-s)^{\alpha-1}\|V(s)\|_2\,ds\Big)^2\,dt \\ + C\int_0^{t_N}\Big(\int_0^t (t-s)^{\alpha-1}\|W(s)\|_2\,ds\Big)^2\,dt.\end{aligned}$$

Here, Young's inequality (Lemma 6.1) and Lemma 3.1 show that the first term is bounded by

$$C\int_0^{t_N} t^{2\alpha-2}\,dt\Big(\int_0^{t_N}\|V\|_2\,dt\Big)^2 \le CT^{2\alpha-1}\tilde{L}_N^2\|\varphi\|^2.$$

For the second term we use Schwarz' inequality to obtain the bound

$$CT^{2\alpha-1}\sum_{n=1}^{N} k_n\int_0^{t_n}\|W\|_2^2\,ds.$$

The proof is again concluded by the standard Gronwall lemma. □

The case of $0 < \alpha \le \frac{1}{2}$ is more involved and requires some preparations. In addition to the technical lemmas of Section 6, we shall need the following simple identity, where we use the piecewise constant function $t \mapsto \hat{t}$ defined by

$$\hat{t} = t_n, \quad \text{for } t \in (t_{n-1}, t_n]. \tag{3.13}$$

Lemma 3.4. *If $\delta \in \mathbf{R}$, $X \in \mathcal{W}_N$, and $Y(t) = \hat{t}^\delta X(t)$, with $\hat{t}$ defined in (3.13), then*

$$\begin{aligned}\sum_{n=1}^{N} \int_{I_n} (X_t(t), Y(t))\, dt &+ \sum_{n=1}^{N-1} ([X]_n, Y_n^+) + (X_0^+, Y_0^+) \\ &= \tfrac12 t_N^\delta \|X_N^-\|^2 + \tfrac12 t_1^\delta \|X_0^+\|^2 + \tfrac12 \sum_{n=1}^{N-1} t_{n+1}^\delta \|[X]_n\|^2 - \tfrac12 \sum_{n=1}^{N-1} \Big(t_{n+1}^\delta - t_n^\delta\Big) \|X_n^-\|^2.\end{aligned}$$

Proof. This follows by a straightforward calculation using $Y_n^+ = t_{n+1}^\delta X_n^+$. □

Lemma 3.5. *Let W be the solution of (3.9). If $B(t,s)$ is weakly singular with $0 < \alpha \le 1$, then we have*

$$\int_0^{t_N} \hat{t}^{-2\alpha} \|W(t)\|_1^2\, dt \le C(T) \|\varphi\|^2.$$

Proof. With $X(t) = \hat{t}^{-2\alpha} W(t)$ in (3.9) we have, in view of Lemma 3.4, the bound for B in (1.2), and Schwarz' inequality,

$$\begin{aligned}&\tfrac12 t_N^{-2\alpha} \|W_N^-\|^2 + \tfrac12 t_1^{-2\alpha} \|W_0^+\|^2 + \tfrac12 \sum_{n=1}^{N-1} t_{n+1}^{-2\alpha} \|[W]_n\|^2 \\ &\quad - \tfrac12 \sum_{n=1}^{N-1} \Big(t_{n+1}^{-2\alpha} - t_n^{-2\alpha}\Big) \|W_n^-\|^2 + \int_0^{t_N} \hat{t}^{-2\alpha} \|W(t)\|_1^2\, dt \\ &\quad = -\int_0^{t_N} \hat{t}^{-2\alpha} \int_0^t B(t,s; U(s), W(t))\, ds\, dt \\ &\quad \le C \int_0^{t_N} \hat{t}^{-2\alpha} \|W(t)\|_1 \int_0^t (t-s)^{\alpha-1} \|U(s)\|_1\, ds\, dt \\ &\quad \le \tfrac12 \int_0^{t_N} \hat{t}^{-2\alpha} \|W(t)\|_1^2\, dt + C \int_0^{t_N} \Big(\hat{t}^{-\alpha} \int_0^t (t-s)^{\alpha-1} \|U(s)\|_1\, ds\Big)^2\, dt.\end{aligned}$$

After deleting nonnegative terms on the left side we hence have

$$\int_0^{t_N} \hat{t}^{-2\alpha} \|W(t)\|_1^2\, dt \le C \int_0^{t_N} \Big(\hat{t}^{-\alpha} \int_0^t (t-s)^{\alpha-1} \|U(s)\|_1\, ds\Big)^2\, dt.$$

Recalling that $U = V + W$ and $t \le \hat{t}$, we have by Lemmas 6.2 ($\beta = \frac12$) and 3.1,

$$\int_0^{t_N} \Big(\hat{t}^{-\alpha} \int_0^t (t-s)^{\alpha-1} \|V(s)\|_1\, ds\Big)^2\, dt \le C \int_0^{t_N} \|V\|_1^2\, dt \le C \|\varphi\|^2.$$

Similarly, since $\hat{s} \le \hat{t}$, Lemma 6.3 (with $\delta = 0$) shows

$$\begin{aligned}\int_0^{t_N} &\Big(\hat{t}^{-\alpha} \int_0^t (t-s)^{\alpha-1} \|W(s)\|_1\, ds\Big)^2\, dt \\ &\le \int_0^{t_N} \Big(\int_0^t (t-s)^{\alpha-1} \big(\hat{s}^{-\alpha} \|W(s)\|_1\big)\, ds\Big)^2\, dt \\ &\le C T^\alpha \int_0^{t_N} (t_N - t)^{\alpha-1} \int_0^t \hat{s}^{-2\alpha} \|W(s)\|_1^2\, ds\, dt.\end{aligned}$$

Replacing t by $\hat{t}$ in the upper limit of the inner integral, we hence have

$$\int_0^{t_N} \hat{t}^{-2\alpha}\|W(t)\|_1^2\,dt \le C\|\varphi\|^2 + C\sum_{n=1}^{N}\Big(\int_{I_n}(t_N-t)^{\alpha-1}\,dt\int_0^{t_n}\hat{s}^{-2\alpha}\|W(s)\|_1^2\,ds\Big),$$

and the proof may be completed by Lemma 6.4. $\square$

Lemma 3.6. *The conclusion of Lemma 3.2 remains valid also if $B(t,s)$ is weakly singular with $0<\alpha\le\frac{1}{2}$.*

Proof. Recall the assumption $k_{n+1}/k_n\le\omega$ for $n\ge1$. This time we shall show

$$\int_0^{t_N}\hat{t}^{2\gamma}\|W(t)\|_2^2\,dt\le C(T,\omega)\tilde{L}_N^2\|\varphi\|^2, \tag{3.14}$$

for $\frac{1}{2}-\alpha<\gamma<\frac{1}{2}$, from which the result obviously follows since $2\gamma<1$. Throughout this proof we use the abbreviation $C=C(T,\omega)$.

With $X(t)=\hat{t}^{2\gamma}AW(t)$ in (3.9) we now have by Lemma 3.4

$$\begin{aligned}\tfrac{1}{2}t_N^{2\gamma}\|W_N^-\|_1^2+\tfrac{1}{2}t_1^{2\gamma}\|W_0^+\|_1^2+\tfrac{1}{2}\sum_{n=1}^{N-1}t_{n+1}^{2\gamma}\|[W]_n\|_1^2+\int_0^{t_N}\hat{t}^{2\gamma}\|W(t)\|_2^2\,dt\\ =\tfrac{1}{2}\sum_{n=1}^{N-1}\Big(t_{n+1}^{2\gamma}-t_n^{2\gamma}\Big)\|W_n^-\|_1^2-\int_0^{t_N}\hat{t}^{2\gamma}\int_0^t B(t,s;U(s),AW(t))\,ds\,dt.\end{aligned}$$

Using the bound for B in (1.2), and Schwarz' inequality, the last term on the right may be estimated by

$$\begin{aligned}C\int_0^{t_N}\hat{t}^{2\gamma}\|W(t)\|_2\int_0^t(t-s)^{\alpha-1}\|U(s)\|_2\,ds\,dt\\ \le\tfrac{1}{2}\int_0^{t_N}\hat{t}^{2\gamma}\|W(t)\|_2^2\,dt+C\int_0^{t_N}\Big(\hat{t}^{\gamma}\int_0^t(t-s)^{\alpha-1}\|U(s)\|_2\,ds\Big)^2\,dt.\end{aligned}$$

Hence, we have

$$\begin{aligned}\int_0^{t_N}\hat{t}^{2\gamma}\|W(t)\|_2^2\,dt\le{}&\sum_{n=1}^{N-1}\Big(t_{n+1}^{2\gamma}-t_n^{2\gamma}\Big)\|W_n^-\|_1^2\\ &+C\int_0^{t_N}\Big(\hat{t}^{\gamma}\int_0^t(t-s)^{\alpha-1}\|V(s)\|_2\,ds\Big)^2\,dt\\ &+C\int_0^{t_N}\Big(\hat{t}^{\gamma}\int_0^t(t-s)^{\alpha-1}\|W(s)\|_2\,ds\Big)^2\,dt\equiv R_1+R_2+R_3.\end{aligned}$$

Here, by the inequality $(1+x)^{2\gamma}-1\le x$ for $x\ge0$, and the assumption $k_{n+1}/k_n\le\omega$,

$$t_{n+1}^{2\gamma}-t_n^{2\gamma}=t_n^{2\gamma}\Big((1+k_{n+1}/t_n)^{2\gamma}-1\Big)\le k_{n+1}t_n^{-1+2\gamma}\le\omega k_nt_n^{-1+2\gamma},$$

so that, by the inverse estimate $k_n|W|_{1,I_n}^2\le C\int_{I_n}\|W\|_1^2\,dt$, and Lemma 3.5,

$$R_1\le\omega\sum_{n=1}^{N-1}k_nt_n^{-1+2\gamma}|W|_{1,I_n}^2\le CT^{2\alpha+2\gamma-1}\int_0^{t_N}\hat{t}^{-2\alpha}\|W(t)\|_1^2\,dt\le C\|\varphi\|^2.$$

We now turn to R_2, and note that

$$\hat{t}^{\gamma}\le C\big((t-s)^{\gamma}+\hat{s}^{1/2}s^{\gamma-1/2}\big),\quad\text{for }0<s<t. \tag{3.15}$$

This inequality clearly holds for $s, t \in I_1$ because $s \le \hat{s} = \hat{t}$ then. For $t \ge t_1$ the inequality is obtained by combination of $\hat{t} \le (1+\omega)t$, for $t \ge t_1$, with the elementary inequality $t^\gamma \le (t-s)^\gamma + s^\gamma$, for $0 < s < t$. In view of (3.15), we thus have

$$R_2 \le C \int_0^{t_N} \Big(\int_0^t (t-s)^{\alpha+\gamma-1} \|V(s)\|_2 \, ds \Big)^2 dt$$
$$+ CT^{2\alpha+2\gamma-1} \int_0^{t_N} \Big(t^{1/2-\alpha-\gamma} \int_0^t (t-s)^{\alpha-1} s^{\gamma-1/2} \big(\hat{s}^{1/2} \|V(s)\|_2 \big) \, ds \Big)^2 dt.$$

Using first Young's inequality (Lemma 6.1) on the first term and Lemma 6.2 (with $\beta = \gamma$) on the second one, and then Lemma 3.1, we get

$$R_2 \le CT^{2\alpha+2\gamma-1} \Bigg(\Big(\int_0^{t_N} \|V\|_2 \, dt \Big)^2 + \int_0^{t_N} \hat{t} \|V(t)\|_2^2 \, dt \Bigg) \le C\tilde{L}_N^2 \|\varphi\|^2.$$

For R_3 we use the inequality $\hat{t}^\gamma \le C\big((t-s)^\gamma s^{-\gamma} \hat{s}^\gamma + \hat{s}^\gamma\big)$, for $0 < s < t$, which is proved in the same way as (3.15). Hence, by Lemma 6.3 (with $\delta = \gamma$ and $\delta = 0$),

$$R_3 \le C \int_0^{t_N} \Big(\int_0^t (t-s)^{\alpha+\gamma-1} s^{-\gamma} \big(\hat{s}^\gamma \|W(s)\|_2 \big) \, ds \Big)^2 dt$$
$$+ C \int_0^{t_N} \Big(\int_0^t (t-s)^{\alpha-1} \big(\hat{s}^\gamma \|W(s)\|_2 \big) \, ds \Big)^2 dt$$
$$\le CT^\alpha \int_0^{t_N} (t_N - t)^{\alpha-1} \int_0^t \hat{s}^{2\gamma} \|W(s)\|_2^2 \, ds \, dt.$$

Replacing t by $\hat{t}$ in the upper limit of the inner integral, and recalling the estimates of R_1 and R_2, we now have, with $C = C(T, \omega)$,

$$\int_0^{t_N} \hat{t}^{2\gamma} \|W(t)\|_2^2 \, dt \le C\tilde{L}_N^2 \|\varphi\|^2 + C \sum_{n=1}^N \int_{I_n} (t_N - t)^{\alpha-1} \int_0^{t_n} \hat{s}^{2\gamma} \|W(s)\|_2^2 \, ds \, dt,$$

and the proof is completed by Lemma 6.4. □

Altogether, from Lemmas 3.1, 3.2, 3.3, and 3.6, we infer that

$$\int_0^{t_N} \|U\|_2 \, dt \le C(T, \omega) \tilde{L}_N \|\varphi\|. \tag{3.16}$$

To complete the proof of Theorem 3.1 it remains to show the following.

Lemma 3.7. *Let U and φ satisfy (3.1). Then*

$$\sum_{n=1}^N \Big(\int_{I_n} \|U_t\| \, dt + \|[U]_{n-1}\| \Big) \le C(T, \omega) \tilde{L}_N \|\varphi\|.$$

Proof. First, (3.6) and (1.2) imply

$$\int_{I_n} (t - t_{n-1}) \|U_t(t)\|^2 \, dt$$
$$\le k_n |U_t|_{I_n} \Big(\int_{I_n} \|U\|_2 \, dt + C \int_{I_n} \int_0^t (t-s)^{\alpha-1} \|U(s)\|_2 \, ds \, dt \Big).$$

In view of (3.7), and $k_n|U_t|_{I_n} \le C\int_{I_n}\|U_t\|\,dt$, we obtain

$$\int_{I_n}\|U_t\|\,dt \le C\Big(\int_{I_n}\|U\|_2\,dt + \int_{I_n}\int_0^t (t-s)^{\alpha-1}\|U(s)\|_2\,ds\,dt\Big), \tag{3.17}$$

and the estimate of U_t follows from (3.16) by a change of order of integration.

Finally, taking $X(t) = [U]_{n-1}$ for $t \in I_n$, $X(t) = 0$ otherwise, in (3.1), we get

$$\begin{aligned}\|[U]_{n-1}\|^2 &= -\int_{I_n}(U_t(t) + AU(t), [U]_{n-1})\,dt - \int_{I_n}\int_0^t B(t,s;U(s),[U]_{n-1})\,ds\,dt\\ &\le \|[U]_{n-1}\|\Big(\int_{I_n}(\|U_t\| + \|U\|_2)\,dt + C\int_{I_n}\int_0^t (t-s)^{\alpha-1}\|U(s)\|_2\,ds\,dt\Big),\end{aligned}$$

and the estimate of $[U]_{n-1}$ follows from (3.16) and (3.17). □

4. The completely discrete method

In this section we consider the completely discrete method for solving (1.1) using the discontinuous Galerkin method for the discretization in time, combined with finite element approximation in space, and prove Theorems 1.5 and 1.6. Thus we now assume $H = L_2(\Omega)$, and that A and $B(t,s)$ are partial differential operators of the form described in Section 1, and that $S_h \subset H_0^1(\Omega)$ satisfies (1.15). The completely discrete solution is now defined by (1.16).

We begin by introducing discrete analogs $A_h, B_h(t,s) : S_h \to S_h$ of $A, B(t,s)$ by

$$(A_h\psi,\chi) = A(\psi,\chi), \quad (B_h(t,s)\psi,\chi) = B(t,s;\psi,\chi), \quad \forall\psi,\chi \in S_h, \tag{4.1}$$

and recall the following analog of (1.2).

Lemma 4.1. *Assume that (1.2) holds and that either (1.18) or (1.19) is satisfied. Then, for A_h and $B_h(t,s)$ defined in (4.1), for $p = 0,1,2,\ p+q = 2$,*

$$|(B_h(t,s)\psi,\chi)| \le C(t-s)^{\alpha-1}\|A_h^{p/2}\psi\|\|A_h^{q/2}\chi\|, \quad \forall\psi,\chi \in S_h.$$

Proof. This is a trivial modification of the proof in [17], pp. 142–143, where the case of a smooth kernel $B(t,s)$ was treated. □

In our analysis of the completely discrete method for (1.1) we introduce for the approximation in the spatial variable the Ritz-Volterra projection defined as the operator $V_h : C([0,T],H^1) \to C([0,T],S_h)$, given by

$$A((V_hu-u)(t),\chi) + \int_0^t B(t,s;(V_hu-u)(s),\chi)\,ds = 0, \quad \forall\chi \in S_h. \tag{4.2}$$

We note that for $t = 0$ (and for all $t \in [0,T]$ if $B(t,s) \equiv 0$), V_hu reduces to the standard Ritz projection defined by

$$A(R_hu-u,\chi) = 0, \quad \forall\chi \in S_h.$$

Lemma 4.2. *With V_hu defined by (4.2) we have, for $t \in [0,T]$,*

$$\|(V_hu)(t) - u(t)\| \le Ch^p\Big(\|u(t)\|_{H^p} + \int_0^t (t-s)^{\alpha-1}\|u(s)\|_{H^p}\,ds\Big), \quad \textit{for } 1 \le p \le r.$$

Proof. This is a minor modification of the proof given in [11], Proposition 2.2. □

Proof of Theorem 1.5. As in the proof of Theorem 1.2 we shall use duality. Let $\|\varphi\| = 1$, and let Z now be the solution of the spatially discrete analog of (1.6) (with $\mathcal{W}_N$ replaced by $\mathcal{W}_{N,h}$). Further, let $W = \Pi V_h u$, where Π is the interpolation operator defined in Section 2. Then we have

$$(U_N^- - u(t_N), \varphi) = (U_N^- - W_N^-, \varphi) + (W_N^- - u(t_N), \varphi) \equiv R + S.$$

Here, using (1.6) and the error equation (1.17),

$$R = G_N(U - W, Z) = (u_{0h} - u_0, Z_0^+) + G_N(u - W, Z) \equiv R_1 + R_2,$$

and hence by (2.2), $[Z]_N = \varphi - Z_N^-$, and the definition of the Ritz-Volterra projection,

$$\begin{aligned} R_2 + S &= \sum_{n=1}^{N} \int_{I_n} \Big(-(u - W, Z_t) + A(u - W, Z) \\ &\quad + \int_0^t B(\cdot, s; (u - W)(s), Z)\, ds \Big)\, dt - \sum_{n=1}^{N} ((u - W)_n^-, [Z]_n) \\ &= \sum_{n=1}^{N} \int_{I_n} \Big(-(u - W, Z_t) + A(V_h u - W, Z) \\ &\quad + \int_0^t B(\cdot, s; (V_h u - W)(s), Z)\, ds \Big)\, dt - \sum_{n=1}^{N} ((u - W)_n^-, [Z]_n). \end{aligned}$$

Application of Lemma 4.1 and Theorem 1.4 (with $H = S_h$, A replaced by A_h, and $\|\cdot\|_2$ replaced by $\|A_h \cdot\|$) now shows

$$\begin{aligned} |R_2 + S| &\le |u - W|_{J_N} \sum_{n=1}^{N} \Big(\int_{I_n} \|Z_t\|\, dt + \|[Z]_n\| \Big) \\ &\quad + C|V_h u - W|_{J_N} \int_0^{t_N} \|A_h Z\|\, dt \\ &\le C L_N \big(|u - W|_{J_N} + |V_h u - W|_{J_N} \big). \end{aligned}$$

Here, in view of the stability of Π (see (2.1) with $m = 0$), we have

$$|u - W|_{J_N} + |V_h u - W|_{J_N} \le C(|(\Pi - I)u|_{J_N} + |V_h u - u|_{J_N}).$$

Using also the obvious bound for R_1 together with (2.1) and Lemma 4.2, this completes the proof of the theorem. $\square$

For the proof of Theorem 1.6 we shall also need the following estimate for the Ritz-Volterra projection, where $\|\cdot\|_2 = \|A \cdot\|$ as before.

Lemma 4.3. *Let $B(t,s)$ be smooth with $B(t,s)$, $B_t(t,s)$, and $B_{tt}(t,s)$ dominated by A, and assume that either (1.18) or (1.19) hold. Then*

$$\|D_t^p A_h (V_h u)(t)\| \le C \Big(\sum_{l=0}^{p} \|D_t^l u(t)\|_2 + \int_0^t \|u\|_2\, ds \Big), \quad \textit{for } p = 0, 1, 2.$$

Proof. For $p = 0$ we have by (4.2), for $\chi \in S_h$,

$$(A_h (V_h u)(t), \chi) = (Au(t), \chi) + \int_0^t B(t, s; (V_h u - u)(s), \chi)\, ds. \tag{4.3}$$

Taking $\chi = A_h(V_h u)(t)$ and using Lemma 4.1 we get

$$\|A_h(V_h u)(t)\| \le \|u(t)\|_2 + C\int_0^t \|A_h V_h u\|\,ds + C\int_0^t \|u\|_2\,ds.$$

Gronwall's lemma then gives

$$\|A_h(V_h u)(t)\| \le C\Big(\|u(t)\|_2 + \int_0^t \|u\|_2\,ds\Big),$$

which is the desired result when $p=0$.

For $p=1$ we have, by differentiating (4.3),

$$\begin{aligned}(D_t A_h(V_h u)(t), \chi) = (D_t Au(t), \chi) &+ B(t,t;(V_h u - u)(t), \chi)\\ &+ \int_0^t B_t(t,s;(V_h u - u)(s), \chi)\,ds,\end{aligned}$$

and the result clearly follows using the domination assumption for $B_t(t,s)$ and its consequences for $B_{h,t}(t,s)$ by Lemma 4.1.

The case $p=2$ obtains by further differentiation and similar arguments. □

Proof of Theorem 1.6. We consider again the representation for R_2+S in the proof of Theorem 1.5. Now we have, since $Z(t) = Z_{n-1}^+ + (t-t_{n-1})Z_t$, and $V_h u - W$ is orthogonal to the constants, cf. the proof of Theorem 1.3,

$$\begin{aligned}\Big|\int_0^{t_N} A(V_h u - W, Z)\,dt\Big| &= \Big|\sum_{n=1}^N \int_{I_n} (A_h(V_h u - W)(t), (t-t_{n-1})Z_t)\,dt\Big|\\ &\le CL_N \max_{1\le n\le N} \big(k_n|(\Pi - I)A_h V_h u|_{I_n}\big).\end{aligned}$$

We also have, again in the same way as in the proof of Theorem 1.3,

$$\Big|\int_0^{t_N}\int_0^t B(t,s;(V_h u - W)(s), Z(t))\,dt\Big| \le CL_N \max_{1\le n\le N} \big(k_n|(\Pi - I)A_h V_h u|_{I_n}\big).$$

The right hand sides of the latter inequalities are now estimated using Lemma 4.3. Also, using the properties of Π,

$$\begin{aligned}&\Big|\sum_{n=1}^N \Big(\int_{I_n} (u - W, Z_t)\,dt + ((u-W)_n^-, [Z]_n)\Big)\Big|\\ &\qquad = \Big|\sum_{n=1}^N \Big(\int_{I_n} (u - V_h u, Z_t)\,dt + ((u-V_h u)_n^-, [Z]_n)\Big)\Big|\\ &\qquad \le CL_N|V_h u - u|_{J_N} \le CL_N h^r \sup_{0\le t\le t_N} \|u(t)\|_{H^r},\end{aligned}$$

and the proof is concluded as in Theorem 1.5. □

Corresponding to the examples given in the introduction we shall now give a few examples in the fully discrete context. We observe first that in the case of a smooth kernel $B(t,s)$ and a smooth solution of (1.1), and with an appropriate choice of u_{0h}, Theorems 1.5 and 1.6 show, for piecewise linear functions ($q=2$), nodal error bounds of orders $O(h^r + k^2)$ and $O(h^r + k^3)$, respectively.

We consider now the case of the homogeneous equation ($f=0$) and initial data u_0 "smooth" but not "higher order compatible" in the sense that $u_0 = 0$ on $\partial\Omega$ but $Au_0 \ne 0$ on $\partial\Omega$. Let thus γ be an arbitrary number with $\gamma < \frac{1}{4}$. Then A^γ does not require boundary conditions, and therefore $u_0 \in D(A^{\gamma+1})$. It then also follows from elliptic regularity that $A^\gamma B(t,s)A^{-\gamma-1}$ is a bounded operator. By Corollary 5.1 below (trivially modified to the case of a smooth kernel), applied to the general differential equation case, we then have

$$\|u(t)\|_{H^{2\gamma+2}} \le C. \tag{4.4}$$

Further, in the smooth kernel case, one may show that then $\|u_{tt}(t)\| \le Ct^{\gamma-1}$ (and $u_t \in C^\gamma(\bar J_N, L_2)$). Thus, from Theorem 1.5, in the smooth kernel case, for $r > 2$, $q = 2$, with u_{0h} appropriately chosen (recall the remark following Theorem 1.3),

$$\|U_N^- - u(t_N)\| \le CL_N\Big(h^{2\gamma+2} + k_1^{1+\gamma} + \max_{2\le n\le N}(k_n^2/t_{n-1}^{1-\gamma})\Big).$$

Similarly, one may show that $\|Au_{tt}(t)\| \le Ct^{\gamma-2}$, so that, from Theorem 1.6, for $r > 2$, $q = 2$,

$$\|U_N^- - u(t_N)\| \le CL_N\Big(h^{2\gamma+2} + k_1^{1+\gamma} + \max_{2\le n\le N}(k_n^3/t_{n-1}^{2-\gamma})\Big).$$

In the case of a weakly singular kernel and $u_0 \in D(A^{1+\gamma})$, $\gamma < \frac14$, the blow-up of u_{tt} as $t^{\alpha-1}$ for small t may interfere with the above result. We still have (4.4) and also, with $u = v + w$, cf. Section 5, $\|w_t(t)\| \le Ct^\alpha$, and $\|w_{tt}(t)\| \le Ct^{\alpha-1}$. It is easily seen that then $w_t \in C^\alpha(\bar J_N, L_2)$. Furthermore, $v_t \in C^\gamma(\bar J_N, L_2)$ and $\|v_{tt}(t)\| \le Ct^{\gamma-1}$ so that $u_t \in C^\beta(\bar J_N, L_2)$, where $\beta = \min(\alpha,\gamma)$ and $\|u_{tt}(t)\| \le Ct^{\beta-1}$. We therefore obtain from Theorem 1.5, in the case $r > 2$, $q = 2$,

$$\|U_N^- - u(t_N)\| \le CL_N\Big(h^{2\gamma+2} + k_1^{1+\beta} + \max_{2\le n\le N}(k_n^2/t_{n-1}^{1-\beta})\Big).$$

5. Some regularity estimates in the weakly singular case

The purpose of this section is to elucidate the regularity properties in the weakly singular case by studying the example of a singular kernel of convolution type, $B(t,s) = (t-s)^{\alpha-1}B$, where B is time independent, in the case of a homogeneous equation, i.e., $f = 0$. Information of the type we are presenting about the behavior of the solution for t near 0 was used in the introduction and Section 4 in discussing the application of our error estimates.

For motivation, we consider first the problem

$$u_t + Au + \int_0^t (t-s)^{\alpha-1}Au(s)\,ds = 0,\ t > 0; \quad u(0) = u_0, \tag{5.1}$$

i.e., the case $B = A$. The analysis may then be carried out by eigenvector expansion, and is reduced to the study of the scalar equation

$$y' + \lambda y + \lambda\int_0^t (t-s)^{\alpha-1}y(s)\,ds = 0,\ t \in (0,1); \quad y(0) = 1, \tag{5.2}$$

where λ is an eigenvalue of A. It may be seen that $y(t)$ is bounded independently of λ and continuous as $t \to 0+$ and hence $|y'(t)| \le C\lambda$. Differentiating, we find that

$$y'' + \lambda y' + \lambda t^{\alpha-1}y(0) + \lambda\int_0^t s^{\alpha-1}y'(t-s)\,ds = 0,$$

and it follows that

$$y''(t) \sim \lambda t^{\alpha-1} + O(\lambda^2), \quad \text{as } t \to 0+. \tag{5.3}$$

Interpreted in terms of (5.1), this means that if $u_0 \in D(A^2)$, then $\|u_{tt}\| = O(t^{\alpha-1})$ as $t \to 0$, and that more regularity of u_0 does not remove this singularity at $t = 0$.

For the purely parabolic problem, i.e., when $B = 0$, the solution is $u(t) = E(t)u_0$, where $E(t) = \exp(-tA)$ is the analytic semigroup generated by $-A$, and

$$\|D_t^k E(t)u_0\|_{2j} \le Ct^{-(k+j-l)}\|u_0\|_{2l}, \quad \text{if } u_0 \in D(A^l),\ 0 \le l \le k+j, \tag{5.4}$$

where l may also take on fractional values. (This estimate is also easy to prove by eigenvector expansion.) In particular, in the differential operator application, the solution has an arbitrary number of derivatives in both x and t for $x \in \Omega$ and $t > 0$. Note that the condition that $u_0 \in D(A^l)$ incorporates various compatibility conditions of u_0 on $\partial\Omega$.

In the case of a weakly singular kernel of the form $B(t,s) = (t-s)^{\alpha-1}B$, the solution $u(t)$ may be written as $u(t) = v(t) + w(t)$, where $v(t) = E(t)u_0$, and where $w(t)$ solves

$$w_t + Aw = -\int_0^t (t-s)^{\alpha-1}Bu(s)\,ds,\ t > 0; \quad w(0) = 0. \tag{5.5}$$

Our aim is now to prove estimates for low order derivatives of w in the case that $u_0 \in D(A)$. Combined with (5.4) these yield estimates for u. More precisely, we shall show the following result which means, in particular, that our present case is no worse, with respect to singular behavior, than the case above when $B = A$.

Theorem 5.1. *If $u_0 \in D(A)$, then we have*

$$\|D_t^k w(t)\|_{2j} \le Ct^{\alpha+1-k-j}\|u_0\|_2, \quad \textit{for } j = 0,1,\ 0 < k+j \le 2. \tag{5.6}$$

We begin with two lemmas.

Lemma 5.1. *Let $g(t)$ be a function with values in H satisfying*

$$\|g(t)\| \le C_0 t^{\alpha-1} \quad \textit{and} \quad \|g'(t)\| \le C_0 t^{\alpha-2}. \tag{5.7}$$

Then, with $E(t)$ the semigroup generated by $-A$, we have

$$\|F(t)\|_2 = \|AF(t)\| \le CC_0 t^{\alpha-1}, \quad \textit{where } F(t) = \int_0^t E(t-s)g(s)\,ds.$$

Proof. We have, since $AE(\sigma) = -E'(\sigma)$,

$$AF(t) = \int_0^{t/2} AE(t-s)g(s)\,ds + \int_{t/2}^t (D_s E(t-s))g(s)\,ds \equiv R_1 + R_2.$$

Here, by (5.4) and (5.7)

$$\|R_1\| \le CC_0 \int_0^{t/2} (t-s)^{-1}s^{\alpha-1}\,ds \le CC_0 t^{\alpha-1}.$$

Integrating by parts,

$$R_2 = g(t) - E(t/2)g(t/2) - \int_{t/2}^t E(t-s)g'(s)\,ds,$$

and the desired result follows. □

Lemma 5.2. *Let $g(t)$ be a function with values in H. Then*

$$\Big\|\int_0^t E(t-s)\int_0^s (s-\sigma)^{\alpha-1}g(\sigma)\,d\sigma\,ds\Big\|_2 \le C\int_0^t (t-\sigma)^{\alpha-1}\|g(\sigma)\|\,d\sigma.$$

Proof. By a switch of order of integration and a change of variables we have

$$\int_0^t E(t-s)\int_0^s (s-\sigma)^{\alpha-1} g(\sigma)\,d\sigma\,ds = \int_0^t \left(\int_0^{t-\sigma} E(t-\sigma-\tau)\tau^{\alpha-1}\,d\tau\right) g(\sigma)\,d\sigma.$$

Application of Lemma 5.1 to the inner integral immediately proves the lemma. □

Proof of Theorem 5.1. We begin with the case $k=0$, $j=1$. Letting $u=v+w$ on the right in (5.5) and using Duhamel's principle,

$$\begin{aligned} w(t) &= -\int_0^t E(t-s)\int_0^s (s-\sigma)^{\alpha-1} Bv(\sigma)\,d\sigma\,ds \\ &\quad -\int_0^t E(t-s)\int_0^s (s-\sigma)^{\alpha-1} Bw(\sigma)\,d\sigma\,ds \equiv R_1+R_2. \end{aligned} \tag{5.8}$$

Writing $B=(BA^{-1})A$, we have from Lemma 5.2 and (1.2) (the assumption that A dominates B) that

$$\|R_2\|_2 \le C\int_0^t (t-\sigma)^{\alpha-1}\|w(\sigma)\|_2 d\sigma.$$

Further, by the boundedness of $E(t)$ we have, again using Lemma 5.2,

$$\begin{aligned} \|R_1\|_2 &= \Big\|\int_0^t AE(t-s)\int_0^s (s-\sigma)^{\alpha-1}(BA^{-1})E(\sigma)Au_0\,d\sigma\,ds\Big\| \\ &\le C\int_0^t (t-\sigma)^{\alpha-1}\|E(\sigma)Au_0\|\,d\sigma \le Ct^\alpha\|u_0\|_2. \end{aligned}$$

Thus

$$\|w(t)\|_2 \le Ct^\alpha\|u_0\|_2 + C\int_0^t (t-\sigma)^{\alpha-1}\|w(\sigma)\|_2\,d\sigma,$$

from which the result follows by Gronwall's lemma, cf. Lemma 1 in [2] and (6.1) below.

We shall next show (5.6) in the case $k=j=1$. Differentiating (5.5) we have

$$w_{tt} + Aw_t = -t^{\alpha-1}Bu_0 - \int_0^t s^{\alpha-1}Bu_t(t-s)\,ds. \tag{5.9}$$

Setting $u=v+w$ and noting that $Aw(0)=0$ by the already proven case of (5.6) and hence $w_t(0)=0$ by (5.5), we have, by Duhamel's principle,

$$\begin{aligned} w_t(t) &= -\int_0^t E(t-s)s^{\alpha-1}Bu_0\,ds \\ &\quad -\int_0^t E(t-s)\int_0^s (s-\sigma)^{\alpha-1}BE_t(\sigma)u_0\,d\sigma\,ds \\ &\quad -\int_0^t E(t-s)\int_0^s (s-\sigma)^{\alpha-1}Bw_t(\sigma)\,d\sigma\,ds \equiv R_1+R_2+R_3. \end{aligned} \tag{5.10}$$

By Lemma 5.1 and (1.2), we find

$$\|R_1\|_2 \le Ct^{\alpha-1}\|u_0\|_2.$$

Further, for R_2 we have

$$\|R_2\|_2 = \Big\|\int_0^t AE(t-s)BA^{-1}g(s)\,ds\Big\|,$$

where, using that $E'(\sigma)u_0 = -E(\sigma)Au_0$,

$$g(t) = \int_0^t (t-s)^{\alpha-1} AE(s)Au_0\,ds = \int_0^t s^{\alpha-1} AE(t-s)Au_0\,ds.$$

We shall apply Lemma 5.1. By that lemma, $\|g(t)\| \le Ct^{\alpha-1}\|u_0\|_2$. For the estimate required for $g'(t)$ we write

$$g(t) = \Big(\int_0^{t/2} + \int_{t/2}^t\Big) s^{\alpha-1} AE(t-s)Au_0\,ds = g_1(t) + g_2(t).$$

Here

$$g_1'(t) = (t/2)^{\alpha-1} AE(t/2)Au_0 - \int_0^{t/2} s^{\alpha-1} A^2E(t-s)Au_0\,ds,$$

and clearly, using (5.4), $\|g_1'(t)\| \le Ct^{\alpha-2}\|u_0\|_2$. Further, by integration by parts and a change of variables, we have

$$\begin{aligned} g_2(t) &= \int_{t/2}^t s^{\alpha-1} D_sE(t-s)Au_0\,ds \\ &= t^{\alpha-1}Au_0 - (t/2)^{\alpha-1}E(t/2)Au_0 - (\alpha-1)\int_0^{t/2} (t-s)^{\alpha-2}E(s)Au_0\,ds. \end{aligned}$$

Differentiating this expression with respect to time, and proceeding as for g_1', gives the same bound for g_2' as for g_1', namely, $\|g_2'(t)\| \le Ct^{\alpha-2}\|u_0\|_2$. We may then apply Lemma 5.1 to obtain

$$\|R_2\|_2 \le Ct^{\alpha-1}\|u_0\|_2.$$

Finally, from Lemma 5.2 and (1.2) we find

$$\|R_3\|_2 \le C\int_0^t (t-\sigma)^{\alpha-1}\|w_t(\sigma)\|_2\,d\sigma.$$

Thus (5.10) gives

$$\|w_t(t)\|_2 \le Ct^{\alpha-1}\|u_0\|_2 + C\int_0^t (t-\sigma)^{\alpha-1}\|w_t(\sigma)\|_2\,d\sigma,$$

and a variant of Gronwall's lemma, cf. Lemma 1 in [2], yields the desired estimate.

We may now also easily treat the case $j = 0$, $k = 2$, i.e., estimate w_{tt} by use of (5.9), splitting the last term on the right there as

$$\int_0^t s^{\alpha-1}Bw_t(t-s)\,ds + \int_0^{t/2} s^{\alpha-1}BE(t-s)Au_0\,ds + \int_{t/2}^t s^{\alpha-1}BE_t(t-s)u_0\,ds.$$

We then apply the result already derived for $j = k = 1$ to estimate the first term and use integration by parts in the last.

The case $k = 1$, $j = 0$ is treated in a similar fashion using (5.5). □

In our discussion at the end of Section 4 we also used the following:

Corollary 5.1. *Let $\gamma < \frac{1}{4}$. If $u_0 \in D(A^{1+\gamma})$ and $\|A^{1+\gamma}BA^{-1-\gamma}\| \le C$, then*

$$\|u(t)\|_{2+2\gamma} \le C\|u_0\|_{2+2\gamma}. \tag{5.11}$$

Proof. This follows by applying $A^{1+\gamma}$ to (5.8), noting that (5.11) holds with $u(t)$ replaced by $v(t)$, and proceeding as before with the appropriate modifications. □

Finally, we present a fact used in the discussion of Theorem 1.3 in the introduction.

Corollary 5.2. *If $B = A$ and $u_0 \in D(A^{2+\alpha})$, then*

$$\|u_{tt}(t)\|_2 \le C t^{\alpha-1} \|u_0\|_{4+2\alpha}.$$

Proof. By the discussion in the beginning of this section it suffices to show that for the solution of (5.2),

$$\lambda |y''(t)| \le C \lambda^{2+\alpha} t^{\alpha-1}.$$

Writing $y = v + w = e^{-\lambda t} + w$ we have by Theorem 5.1 (with $H = R, B = A =$ multiplication by λ) that $|w''(t)| \le \lambda t^{\alpha-1}$. The appropriate bound for v follows from (5.4). □

6. Some technical lemmas

In this section we state and prove three inequalities and a variant of Gronwall's lemma, that we have used several times above.

We begin with the following variant of Young's inequality, which follows at once by the standard Young's inequality for convolutions on $\mathbf{R}$ by extending the functions by 0 outside the interval $[0, T]$.

Lemma 6.1. *If $f \in L_1([0,T])$ and $g \in L_2([0,T])$, then*

$$\Big(\int_0^T \Big(\int_0^t f(t-s) g(s)\, ds \Big)^2 dt \Big)^{1/2} \le \int_0^T |f(t)|\, dt \Big(\int_0^T g^2(t)\, dt \Big)^{1/2}.$$

We continue with a special case of an inequality of Hardy, Littlewood and Pólya, adapted from Theorem (6.20), p. 187, in [6].

Lemma 6.2. *If $\alpha, \beta > 0$, $\kappa = \big(\int_0^1 (1-t)^{\alpha-1} t^{\beta-1}\, dt \big)^2$, and $f \in L_2([0,T])$, then*

$$\int_0^T \Big(t^{1/2-\alpha-\beta} \int_0^t (t-s)^{\alpha-1} s^{\beta-1/2} f(s)\, ds \Big)^2 dt \le \kappa \int_0^T f^2(t)\, dt.$$

Proof. With $\sigma = s/t$ we have

$$t^{1/2-\alpha-\beta} \int_0^t (t-s)^{\alpha-1} s^{\beta-1/2} f(s)\, ds = \int_0^1 (1-\sigma)^{\alpha-1} \sigma^{\beta-1/2} f(t\sigma)\, d\sigma,$$

and, by Minkowski's inequality for integrals (see [6], p. 186),

$$\begin{aligned} \Big(\int_0^T \Big(\int_0^1 & (1-\sigma)^{\alpha-1} \sigma^{\beta-1/2} f(t\sigma)\, d\sigma \Big)^2 dt \Big)^{1/2} \\ &\le \int_0^1 \Big(\int_0^T \big((1-\sigma)^{\alpha-1} \sigma^{\beta-1/2} f(t\sigma) \big)^2 dt \Big)^{1/2} d\sigma \\ &= \int_0^1 (1-\sigma)^{\alpha-1} \sigma^{\beta-1/2} \Big(\int_0^T f^2(t\sigma)\, dt \Big)^{1/2} d\sigma. \end{aligned}$$

Since $\int_0^T f^2(t\sigma)\, dt \le \sigma^{-1} \int_0^T f^2(t)\, dt$ for $0 < \sigma < 1$, this proves the desired inequality. □

The following is a generalization of Lemma 2 in [2].

Lemma 6.3. *If $\alpha > 0$, $0 \le \delta < \frac{1}{2}$, $\mu = \int_0^1 (1-t)^{\alpha+\delta-1} t^{-2\delta}\, dt$, and $f \in L_2([0,T])$, then*

$$\int_0^T \Big(\int_0^t (t-s)^{\alpha+\delta-1} s^{-\delta} f(s)\, ds \Big)^2 dt \le \mu T^\alpha \int_0^T (T-t)^{\alpha-1} \int_0^t f^2(s)\, ds\, dt.$$

Proof. By Schwarz' inequality we have

$$\Big(\int_0^t (t-s)^{\alpha+\delta-1} s^{-\delta} f(s)\, ds \Big)^2 \le \int_0^t (t-s)^{\alpha+\delta-1} s^{-2\delta}\, ds \int_0^t (t-s)^{\alpha+\delta-1} f^2(s)\, ds,$$

where, for $0 \le s < t \le T$,

$$\int_0^t (t-\tau)^{\alpha+\delta-1} \tau^{-2\delta}\, d\tau = \mu t^{\alpha-\delta} \le \mu T^\alpha (t-s)^{-\delta}.$$

Hence

$$\int_0^T \Big(\int_0^t (t-s)^{\alpha+\delta-1} s^{-\delta} f(s)\, ds \Big)^2 dt \le \mu T^\alpha \int_0^T \int_0^t (t-s)^{\alpha-1} f^2(s)\, ds\, dt.$$

The integral on the right side is equal to

$$\begin{aligned} \int_0^T \int_0^t s^{\alpha-1} f^2(t-s)\, ds\, dt &= \int_0^T s^{\alpha-1} \int_s^T f^2(t-s)\, dt\, ds \\ &= \int_0^T s^{\alpha-1} \int_0^{T-s} f^2(\tau)\, d\tau\, ds = \int_0^T (T-t)^{\alpha-1} \int_0^t f^2(\tau)\, d\tau\, dt, \end{aligned}$$

which yields the desired result. □

A well known version of Gronwall's lemma states that if ϕ and a are nonnegative functions with a increasing, then

$$\begin{aligned} &\phi(t) \le a(t) + K \int_0^t (t-s)^{\alpha-1} \phi(s)\, ds, \quad \text{for } t \in [0,T], \\ &\text{implies} \quad \phi(t) \le C(K,\alpha,T) a(t), \quad \text{for } t \in [0,T], \end{aligned} \tag{6.1}$$

cf. Lemma 1 in [2]. The following is a discrete version of this, that we have used repeatedly in our treatment of the case of a weakly singular kernel. We remind the reader that $k = \max_n k_n$.

Lemma 6.4. *Assume that $0 < \alpha \le 1$, $K \ge 0$, $\phi_N, a_N \ge 0$, $a_N \le a_{N+1}$ for $N \ge 1$, and $\delta = K k^\alpha / \alpha < 1$. If, for $t_N \in (0,T]$,*

$$\phi_N \le a_N + K \sum_{n=1}^N \omega_{N,n}^{(\alpha)} \phi_n, \quad \textit{where } \omega_{N,n}^{(\alpha)} = \int_{I_n} (t_N - t)^{\alpha-1}\, dt, \tag{6.2}$$

then

$$\phi_N \le C(\delta, K, \alpha, T) a_N, \quad \textit{for } t_N \in (0,T].$$

Proof. We begin by showing that, for $0 < \alpha \le 1$, $\beta > 0$, we have

$$\sum_{n=j+1}^N \omega_{N,n}^{(\beta)} \omega_{n,j}^{(\alpha)} \le C \omega_{N,j}^{(\beta+\alpha)}, \quad j \le N-1, \tag{6.3}$$

where $C = \int_0^1 (1-t)^{\beta-1} t^{\alpha-1}\,dt$. This follows from the following calculation, where we employ the inequalities $(t_n - s)^{\alpha-1} \le (t-s)^{\alpha-1}$ and $s \le t_j$:

$$\begin{aligned}
\sum_{n=j+1}^{N} \omega_{N,n}^{(\beta)} \omega_{n,j}^{(\alpha)} &\le \sum_{n=j+1}^{N} \int_{I_j} \int_{I_n} (t_N - t)^{\beta-1} (t-s)^{\alpha-1}\,dt\,ds \\
&= \int_{I_j} \Big(\int_{t_j}^{t_N} (t_N - t)^{\beta-1} (t-s)^{\alpha-1}\,dt \Big)\,ds \\
&\le \int_{I_j} \Big(\int_{s}^{t_N} (t_N - t)^{\beta-1} (t-s)^{\alpha-1}\,dt \Big)\,ds \\
&= C \int_{I_j} (t_N - s)^{\beta+\alpha-1}\,ds = C \omega_{N,j}^{(\beta+\alpha)}.
\end{aligned}$$

Next we note that, in view of the given inequality (6.2),

$$\phi_N \le a_N + K \sum_{n=1}^{N-1} \omega_{N,n}^{(\alpha)} \phi_n + K \omega_{N,N}^{(\alpha)} \phi_N, \quad N \ge 1.$$

Since $K\omega_{N,N}^{(\alpha)} = K k_N^{\alpha}/\alpha \le \delta$, we get

$$\phi_N \le (1-\delta)^{-1} \Big(a_N + K \sum_{n=1}^{N-1} \omega_{N,n}^{(\alpha)} \phi_n \Big), \quad N \ge 1. \tag{6.4}$$

Then we prove, by induction, that, for $l = 1, 2, \dots,$

$$\phi_N \le C(\delta, K, l, \alpha, T) \Big(a_N + \sum_{n=1}^{N-1} \omega_{N,n}^{(l\alpha)} \phi_n \Big), \quad t_N \in (0, T]. \tag{6.5}$$

This is clearly true for $l = 1$. For the induction step we use (6.4) to get

$$\sum_{n=1}^{N-1} \omega_{N,n}^{((l-1)\alpha)} \phi_n \le C(\delta, K, l, \alpha, T) \Big(\sum_{n=1}^{N-1} \omega_{N,n}^{((l-1)\alpha)} a_n + \sum_{n=1}^{N-1} \omega_{N,n}^{((l-1)\alpha)} \sum_{j=1}^{n-1} \omega_{n,j}^{(\alpha)} \phi_j \Big),$$

where, to estimate the first term, we note that

$$\sum_{n=1}^{N-1} \omega_{N,n}^{((l-1)\alpha)} a_n \le \Big(\sum_{n=1}^{N-1} \omega_{N,n}^{((l-1)\alpha)} \Big) a_N = C(l, \alpha, T) a_N,$$

and, by switching the order of summation and using (6.3),

$$\sum_{n=1}^{N-1} \omega_{N,n}^{((l-1)\alpha)} \sum_{j=1}^{n-1} \omega_{n,j}^{(\alpha)} \phi_j = \sum_{j=1}^{N-2} \Big(\sum_{n=j+1}^{N-1} \omega_{N,n}^{((l-1)\alpha)} \omega_{n,j}^{(\alpha)} \Big) \phi_j \le C(l, \alpha) \sum_{j=1}^{N-1} \omega_{N,j}^{(l\alpha)} \phi_j.$$

This proves (6.5).

Finally, we choose l such that $l\alpha - 1 \ge 0$. Then $\omega_{N,n}^{(l\alpha)} = \int_{I_n} (t_N - t)^{l\alpha-1}\,dt \le T^{l\alpha-1} k_n$, so that (6.5) implies

$$\phi_N \le C(\delta, K, \alpha, T) \Big(a_N + \sum_{n=1}^{N-1} k_n \phi_n \Big), \quad t_N \in (0, T].$$

The result now immediately follows by the standard discrete Gronwall lemma. □

Acknowledgements. We thank an anonymous referee and James H. Bramble for several suggestions leading to a more lucid presentation of our results.

References

[1] J. R. Cannon and Y. Lin, *Nonclassical H^1 projection and Galerkin methods for nonlinear parabolic integrodifferential equations*, Calcolo **25** (1988), 187–201.

[2] C.-M. Chen, V. Thomée, and L. B. Wahlbin, *Finite element approximation of a parabolic integro-differential equation with a weakly singular kernel*, Math. Comp. **58** (1992), 587–602.

[3] B. R. Coleman and M. E. Gurtin, *Equipresence and constitutive equations for rigid heat conductors*, Z. Angew. Math. **18** (1967), 199–208.

[4] 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.

[5] 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.

[6] G. B. Folland, *Real Analysis. Modern Techniques and Their Applications*, John Wiley and Sons, New York, 1984.

[7] R. C. Grimmer and A. J. Pritchard, *Analytic resolvent operators for integral equations in Banach space*, J. Differential Equations **50**, (1983), 234–259.

[8] Y. Lin, V. Thomée, and L. B. Wahlbin, *Ritz-Volterra projections to finite-element spaces and applications to integrodifferential and related equations*, SIAM J. Numer. Anal. **28** (1991), 1047–1070.

[9] A. Lorenzi and E. Sinestrari, *An inverse problem in the theory of materials with memory*, Nonlinear Anal. Theory Methods Appl. **12**, (1988), 1317–1335.

[10] A. Lunardi and E. Sinestrari, *C^α-regularity of nonautonomous linear integro-differential equations of parabolic type*, J. Differential Equations **63**, (1986), 88–116.

[11] A. K. Pani, V. Thomée, and L. B. Wahlbin, *Numerical methods for hyperbolic and parabolic integro-differential equations*, J. Integral Equations Appl. **4** (1992), 533–584.

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

[13] V. Thomée and L. B. Wahlbin, *Long time numerical solution of a parabolic equation with memory*, Math. Comp. **62** (1994), 477–496.

[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] ______, *Backward Euler type methods for parabolic integro-differential equations with nonsmooth data*, Contributions in Numerical Mathematics (Singapore) (R. P. Agarwal, ed.), World Scientific Series in Applicable Analysis, vol. 2, World Scientific Publishing Co., Singapore, 1993, pp. 373–388.

[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 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 and Göteborg University, S–412 96 Göteborg, Sweden
E-mail address: `stig@math.chalmers.se`

Department of Mathematics, Chalmers University of Technology and Göteborg University, S–412 96 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`