FINITE ELEMENT APPROXIMATION OF DIFFUSION EQUATIONS WITH CONVOLUTION TERMS

MAŁGORZATA PESZYŃSKA

ABSTRACT. Approximation of solutions to diffusion equations with memory represented by convolution integral terms is considered. Such problems arise from modeling of flows in fissured media. Convergence of the method is proved and results of numerical experiments confirming the theoretical results are presented. The advantages of implementation of the algorithm in a multiprocessing environment are discussed.

1. INTRODUCTION

We present a convergent method of fully-discrete approximation of solutions of diffusion equations with integral convolution terms. Equations of this type arise in the modeling of evolution phenomena where memory effects occur. In our research we were motivated by a problem of diffusion in fissured media described in [14] (see also [3, 25]) in the form

$$(1) \qquad u_t + Lu_t - \nabla \cdot (\mathbf{D}\nabla u) = f(x,t) \ , \quad (x,t) \in Q,$$

where u is the unknown density or concentration of the fluid, $u_t = \frac{\partial u}{\partial t}$, $\mathbf{D}$ is the diffusivity tensor and f is the external source term. L is the convolution operator

$$Lv = \tau * v = \int_0^t \tau(s)v(t-s)ds$$

with kernel τ which is a continuous real–valued integrable function. Its values are derived from microscopic properties of the domain of the flow; see [14, 25]. The equation is to be solved in the cylinder $Q = \Omega \times I$ with time interval $I = (0,T)$ and open bounded spatial domain $\Omega \subset \mathbf{R}^d, d = 2,3$, with smooth boundary $\Gamma = \partial\Omega$. The homogeneous Dirichlet boundary condition and the initial condition

$$(2) \qquad u(x,t) \ = \ 0 \ , \quad x \in \Gamma \ , \ t \in I \ ,$$

$$(3) \qquad u(x,0) \ = \ u_0(x) \ , \quad x \in \Omega,$$

complete the model.

Note that the equation (1) for $\tau \equiv 0$ (or τ equal to the Dirac distribution) reduces to the standard diffusion equation. The diffusion equation is derived on the assumption that the response of the solutions to the control components of the

model is "instantaneous." It is well known that real-world phenomena are rarely instantaneous, and it is important to consider methods of analysis and approximation applicable to problems where memory effects occur.

In modeling of flows in fissured media one has to take into account the "history of the flow". In *distributed microstructure models of flows in fissured media* (for modeling, analysis and approximation see [1, 30]) the history is modeled by coupling with the microstructure, and the model has a structure different from (1). (For the discussion of both models see [2, 25].) The memory effects in (1) are represented by the convolution integral term Lu_t. The graph of the convolution kernel τ met in applications lies usually between the two extremes, i.e., "between" $\tau \equiv 0$ and Dirac distribution; τ is positive, integrable and decreasing, in general unbounded with a singularity at the origin. The character of τ has an essential influence on the solution to the model. (See the discussion on patterns of flow corresponding to different convolution kernels in [26].) The possible singularity of τ makes the analysis and approximation of solutions a delicate problem.

Let $\tau \in L^1(I)$. Then obviously the linear operator L is bounded on $L^p(I), p \geq 1$. Moreover, if τ is a *positive* kernel then L is *monotone* on $L^2(I)$ (see [12, 14, 18]). This property leads to the existence-uniqueness results for the problem (1)–(3) (see [14, 25, 28]). Our goal here is to present a method of numerical approximation of solutions to (1). We assume that

$$(4) \qquad \tau \in L^1(I) \cap C(I) , \quad \tau(t) \geq 0, \forall t \in I,$$

$$\tau(\cdot) \text{ is a monotone nonincreasing function.}$$

Our assumptions on τ are slightly weaker than those required for the well–posedness of the model.

We note that memory integral terms arise in different applications (e.g., theory of heat conduction with memory [21, 23] and homogenization limits of conservation laws [19, 32, 33]), but not much research has been devoted to their approximation. On the other hand, approximation of Volterra integrals of the type arising in viscoelasticity was studied for smooth and bounded kernels in [6, 15, 22, 24, 31] and recently in [20, 34, 36] for more general data. Further, the author of [17] studies numerical approximation of convolution integrals in the framework of discretized operational calculus; these results however were unknown to us at the time our research was done. In the framework of *product integration methods* (see [16]), which we adopt here, general approximation methods for integral terms may be constructed. In the approximation method presented below we combine finite elements for discretization in space with a counterpart of the implicit Euler scheme for discretization in time, with special attention applied to the integral convolution terms. Except for that treatment, the technique applied below is standard for finite element approximation of diffusion problems (see [29, 35]). In order to decrease the overall computational time spent for the calculations, we suggest implementing the numerical algorithm in a multiprocessing environment. We address also the issue of computer memory storage which is critical for the implementation. We note that the algorithm presented in this paper provides a basis on which approximation to solutions of complicated models of flows can be constructed (see [26, 27]).

The plan of the paper is as follows. In Section 2 we define the approximation to the solutions of (1)–(3) and prove its stability. In Section 3 the main result, i.e., the convergence of the fully discrete scheme, is proved. Section 4 contains results of

numerical experiments which confirm the convergence theorem. In the last section we suggest implementation ideas that improve effectiveness of the algorithm.

Throughout the paper we use standard notation and results on Sobolev spaces and finite element approximation. For the details we refer, e.g., to [7]. We set $V = H_0^1(\Omega)$, $H = L^2(\Omega)$, and denote by $(\cdot, \cdot)$ the scalar product in H, by $\| \cdot \|$ the norm in V, by $|\cdot|$ the norm in H, and by c_P the constant in the Poincaré inequality $|v| \leq c_P \| v \|, v \in V$. We use $\| \cdot \|_{k,m,t}$ to denote $\| \cdot \|_{L^k((0,t);H^m(\Omega))}$.

We also refer frequently to the function $\mathcal{T}(t) \stackrel{\text{def}}{=} \int_0^t \tau(s)ds = \| \tau \|_{L^1(0,t)}, t \in \bar{I}$, and we denote $C_\tau(t) = \mathcal{T}(t) + 1$. Also, C denotes a generic positive constant, independent of essential quantities (in particular, of discretization parameters).

2. DEFINITION OF APPROXIMATION

In this section we define an approximation scheme for diffusion equations with convolution terms. We first pose an abstract Cauchy problem corresponding to (1)–(3) and prove the related stability result. Then we specify the discretization with respect to the time variable and the appropriate discrete counterpart of the convolution terms and prove stability of the defined semidiscrete scheme.

In what follows we restrict ourselves for simplicity to plane convex polygonal domains $\Omega \subset \mathbf{R}^2$ and to constant symmetric diffusivity tensors $\mathbf{D}$ satisfying the *ellipticity* condition

(5)

$$\sum_{i,j=1}^2 D_{i,j}\xi_i\xi_j \geq d_0 \sum_{i=1}^2 \xi_i^2 \ , \ \ \forall \xi = [\xi_1, \xi_2]^T \in \mathbf{R}^2 \ , \ \ D_{i,j} = D_{j,i}, 1 \leq i,j \leq 2 \ .$$

We also assume

(6) $$f \in L^2(\mathcal{Q}) \ , \qquad u_0 \in H^2(\Omega) \cap V.$$

The abstract Cauchy problem corresponding to (1)–(3) is

(7) $$u_t + Lu_t + \mathbf{M}u \ = \ f(t) \ , \ \ t > 0$$

(8) $$u(0) \ = \ u_0$$

with linear operator $\mathbf{M} = -\nabla \cdot \mathbf{D}\nabla$ defined on the domain $D(\mathbf{M}) = H^2(\Omega) \cap V$ dense in H. The operator $\mathbf{M}$ is unbounded in H, but with the above assumptions on $\mathbf{D}$ is positive definite, and it is an isomorphism from $D(\mathbf{M})$ to H. The bilinear form associated with $\mathbf{M}$ is given by

$$a(u,v) = \sum_{i,j=1}^2 D_{i,j} \int_\Omega \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} dx \ , \ \ u,v \in V$$

and by (5) is symmetric, continuous and V–elliptic, i.e., the following inequalities hold:

$$\begin{aligned} a(u,u) &\geq c_a \| u \|^2 \ , \ \ \forall u \in V \ , \\ a(u,v) &\leq k_a \| u \| \| v \| \ , \ \ \forall u,v \in V \ , \end{aligned}$$

with positive constants c_a, k_a. Also, for some constant $c_D > 0$ we have

$$a(u,v) = (\mathbf{M}u, v) \leq c_D \| u \|_{H^2(\Omega)} |v| \ , \ \ u \in D(\mathbf{M}), v \in V.$$

Before proceeding to define the discrete counterpart of the problem (1)–(3) , we first prove the stability result for the continuous case. This result is given only

for purposes of exposition, as its proof should make it easier to understand the corresponding results for the discrete case given later. For the sake of exposition we restrict ourselves momentarily to *bounded* convolution kernels, an assumption which is by no means necessary in our main results below. The continuous stability result for the general case of unbounded kernels is proven in the Appendix to this paper.

Lemma 1. *Let τ satisfy (4) and let additionally τ be differentiable and defined at $t = 0$. Let also u be a solution of (7)–(8). Then u satisfies*

$$|u(t)| \leq C_\tau(t)|u_0| + \| f \|_{1,0,t} .$$

Proof. Equation (7) can be rewritten by using properties of convolution and the fact that $\tau(0)$ exists (the singular case is treated in the Appendix),

$$(9) \qquad u_t + \tau' * u + \tau(0)u(t) - \tau(t)u(0) + \mathbf{M}u = f.$$

Multiply (9) by u and integrate the resulting identity over Ω to obtain

$$\frac{d}{dt}|u||u(t)| + \big((\tau' * u)(t), u(t)\big) + \tau(0)|u(t)|^2$$
$$- \tau(t)\big(u(0), u(t)\big) + \big(\mathbf{M}u(t), u(t)\big) = \big(f(t), u(t)\big) \ , \ t > 0.$$

Now use Schwarz inequality, positive definiteness of $\mathbf{M}$ and the fact that $\tau'(t) \leq 0$ to estimate

$$(10) \quad \frac{d}{dt}|u||u(t)| + \tau(0)|u(t)|^2$$
$$\leq \tau(t)|u(0)||u(t))| + |f(t)||u(t)| + \big((-\tau') * |u|\big)(t)|u(t)| \ , \ t > 0.$$

Consider the set $J = \{t \in I : |u(t)| = 0\}$. At those $t \in I - J$, we have

$$(11) \qquad \frac{d}{dt}|u| + \tau(0)|u(t)| \leq \tau(t)|u(0)| + |f(t)| + \big((-\tau') * |u|\big)(t) \ , \ t > 0.$$

Since the closed set J consists of accumulation points (at which the inequality (11) holds because the left side is zero) and a countable number of isolated points, we may integrate over $(0, t)$ to obtain

$$(12) \quad |u(t)| + \tau(0)\int_0^t |u(s)|ds \leq |u(0)| \left(1 + \int_0^t \tau(s)ds\right) + \int_0^t |f(s)|ds$$
$$+ \tau(0)\int_0^t |u(s)|ds - (\tau * |u|)(t) \ , t > 0 .$$

The last two terms arise from a change of order of integration in the last term of the previous identity. Finally, positivity of τ and (8) yield

$$|u(t)| \leq |u_0|(1 + \mathcal{T}(t)) + \int_0^t |f(s)|ds = C_\tau(t)|u_0| + \| f \|_{1,0,t}$$

as desired.   □

Next we define the discretization in time. Take an integer $n > 0$, define the time step $\Delta t = T/n$ and split I into n subintervals $I_k, 1 \leq k \leq n$, of length Δt, with $t_k = k\Delta t$. We deal with approximations of functions defined on I. The piecewise linear approximation of a function v defined in I is the function $\bar{v}$ interpolating the discrete values $v(t_k), 0 \leq k \leq n$, hence, linear in each I_k. The piecewise

constant approximation $\widehat{v}$ is defined by the set of mean values $\widehat{v}_k = \frac{1}{\Delta t} \int_{I_k} v(s)ds$ as $\widehat{v}(t) = \sum_{k=1}^{n} \widehat{v}_k \chi_k(t)$, where $\chi_k(\cdot)$ denotes the characteristic function of I_k.

Now we specify the discrete counterpart for the convolution term. Let $v \in L^2(I)$ and let τ satisfy (4). The piecewise constant approximation to Lv is defined by the set of values

$$\widehat{Lv}\,|_{I_j} \overset{\text{def}}{=} \frac{1}{\Delta t} \int_{I_j} (Lv)(s)ds, \quad 1 \le j \le n.$$

We set

$$\theta_{j,k} \overset{\text{def}}{=} \frac{1}{\Delta t} \int_{I_j} \int_0^t \tau(t-r)\chi_k(r)dr dt, \quad 1 \le j, k \le n,$$

so that in the particular case when v is a step function, i.e., $v(t) = \sum_{k=1}^{n} v_k \chi_k(t)$, the piecewise constant approximation $\widehat{Lv}$ of Lv reads

$$\widehat{Lv}|_{I_j} = \sum_{k=1}^{j} \theta_{j,k} v_k.$$

Define

$$\eta_r \overset{\text{def}}{=} \theta_{r+1,1}, \quad 0 \le r < n.$$

Note that $\theta_{j,k} = \theta_{j-1,k-1} = \cdots = \theta_{j-k+1,1}, j \ge k$, hence, $\theta_{j,k} = \eta_{j-k}$, so for a step function v we have

$$\widehat{Lv}|_{I_j} = \sum_{k=1}^{j} \eta_{j-k} v_k.$$

This last representation is very convenient, because it displays the convolution character of the discrete counterpart of L. Below we establish some auxiliary technical facts relevant to it.

Lemma 2. *Let τ satisfy* (4). *Then the coefficients* $(\eta_r)_{r=0}^{n-1}$ *have the following properties:*

  (i) $\eta_r \ge 0, \quad 0 \le r < n;$
  (ii) $\sum_{r=0}^{j-1} \eta_r = \sum_{k=1}^{j} \eta_{j-k} \le \mathcal{T}(t_j), \quad 1 \le j \le n.$
  (iii) *Define* $\vartheta_r \overset{\text{def}}{=} \eta_{r-1} - \eta_r, \quad 1 \le r < n.$ *Then* $\vartheta_r \ge 0$ *for* $r \ge 2$, *and for* Δt *sufficiently small,* $1 + \vartheta_1 \ge 0$. *Moreover,*

$$(13) \qquad 1 + \sum_{k=1}^{j-1} \vartheta_{j-k} \le 1 + \eta_0, \quad 2 \le j \le n.$$

Proof. The first property follows directly from nonnegativity of τ and the definition of $\theta_{j,k}$, which implies $\theta_{j,k} \ge 0$ for $j \ge k$ and $\theta_{j,k} = 0$ otherwise.

Note that $\mathcal{T}(\cdot)$ is continuous, positive and nondecreasing. By definition we have

$$\sum_{k=1}^{j} \eta_{j-k} = \sum_{k=1}^{j} \theta_{j,k} = \frac{1}{\Delta t} \int_{I_j} \int_0^t \tau(t-r) \sum_{k=1}^{j} \chi_k(r) dr dt$$

$$= \frac{1}{\Delta t} \int_{I_j} \int_0^t \tau(r) dr dt = \frac{1}{\Delta t} \int_{I_j} \mathcal{T}(t) dt \le \mathcal{T}(t_j),$$

and this implies (ii). Further, since for $r \ge 2$ the values $\theta_{r,k}$ decrease as k decreases, we obtain $\vartheta_r = \theta_{r,1} - \theta_{r+1,1} = \theta_{r+1,2} - \theta_{r+1,1} \ge 0$. To obtain $1 + \vartheta_1 = 1 + \eta_0 - \eta_1 \ge 0$, it is enough to show $1 \ge \eta_1$ or $\eta_0 \ge \eta_1$. While the latter is satisfied by most functions

met in applications, the former (which we will use) holds for all functions satisfying (4) and Δt sufficiently small. This follows because $\mathcal{T}(t) \to 0$ as $t \to 0$ and

$$\eta_1 = \frac{1}{\Delta t} \int_0^{\Delta t} (\mathcal{T}(t + \Delta t) - \mathcal{T}(t)) dt \leq \mathcal{T}(2\Delta t) \leq 1,$$

if Δt is close to 0.

To complete the proof of the lemma, we calculate

$$1 + \sum_{k=1}^{j-1} \vartheta_{j-k} = 1 + \eta_0 - \eta_{j-1} \leq 1 + \eta_0,$$

and so (iii) follows.   □

The above construction leads to the discrete–in–time scheme for (7)–(8). Define $u_1, u_2, \ldots, u_n$ as solutions of

$$(14) \qquad \partial_n u_j + \sum_{k=1}^{j} \eta_{j-k} \partial_n u_k + \mathbf{M} u_j = \widehat{f_j} \ , \ \ 1 \leq j \leq n \ ,$$

where $\partial_n u_k$ denotes the quotient $\partial_n u_k = \frac{u_k - u_{k-1}}{\Delta t}, 1 \leq k \leq n$, the term u_0 is the prescribed initial data, and $\widehat{f_j}$ corresponds to the piecewise constant approximation of f defined above. Below we prove that this semidiscrete scheme is stable, and in the next section we show convergence of its fully discrete counterpart. Both of these results rely on similar ideas, applicable to discrete equations of the form similar to (14), with general right-hand side terms. We note also the correspondence between the discrete results and the proof for the continuous case given in Lemma 1.

The following auxiliary result is a counterpart of the discrete Gronwall lemma.

Lemma 3. *For arbitrary sequences* $(a_k)_{k=1}^{n}, (b_k)_{k=1}^{n}$ *of positive numbers such that*

$$a_j \leq b_j + \max_{1 \leq k \leq j-1} a_k \ , \ \ 1 \leq k \leq n \ ,$$

we have

$$a_i \leq \sum_{j=1}^{i} b_j, \quad 1 \leq i \leq n.$$

Lemma 4. *Let* τ *satisfy* (4) *and* Δt *be small. Then the solution of the discrete equation*

$$(15) \qquad \partial_n v_j + \sum_{k=1}^{j} \eta_{j-k} \partial_n v_k + \mathbf{M} v_j = F_j, \quad 1 \leq j \leq n,$$

satisfies the estimate

$$|v_i| \leq C_\tau(t_i)|v_0| + \sum_{j=1}^{i} \Delta t |F_j|, \quad 1 \leq i \leq n \ .$$

Proof. Rewrite (15) to get (compare with (9))

$$\frac{v_j - v_{j-1}}{\Delta t} - \sum_{k=1}^{j-1} \frac{\eta_{j-k-1} - \eta_{j-k}}{\Delta t} v_k + \frac{\eta_0}{\Delta t} v_j - \frac{\eta_{j-1}}{\Delta t} v_0 + \mathbf{M} u_j = F_j$$

or, in the weak form,

$$(16) \quad \left(\frac{v_j - v_{j-1}}{\Delta t}, v\right) - \sum_{k=1}^{j-1} \frac{\eta_{j-k-1} - \eta_{j-k}}{\Delta t} (v_k, v) + \frac{\eta_0}{\Delta t} (v_j, v) - \frac{\eta_{j-1}}{\Delta t} (v_0, v)$$

$$+ (\mathbf{M}v_j, v) = (F_j, v) \quad , \quad v \in H.$$

Since ϑ_{j-k} is nonnegative for $j - k \geq 2$ and $1 + \vartheta_1 \geq 0$ (by Lemma 2.iii), setting $v = v_j$ and using the Schwarz inequality and positivity of $\mathbf{M}$ we obtain (see (10))

$$\frac{|v_j|^2}{\Delta t} + \frac{\eta_0}{\Delta t}|v_j|^2 \leq \frac{\eta_{j-1}}{\Delta t}|v_0||v_j| + |F_j||v_j| + (1 + \vartheta_1)\frac{|v_{j-1}|}{\Delta t}|v_j| + \sum_{k=1}^{j-2} \frac{\vartheta_{j-k}}{\Delta t}|v_k||v_j| .$$

Divide both sides by $|v_j|$, multiply by Δt, and apply Lemma 2.iii to get (compare (12))

$$(1 + \eta_0)|v_j| \leq \eta_{j-1}|v_0| + \Delta t|F_j| + (1 + \vartheta_1)|v_{j-1}| + \sum_{k=1}^{j-2} \vartheta_{j-k} \max_{1 \leq k \leq j-2} |v_k|$$

$$\leq \eta_{j-1}|v_0| + \Delta t|F_j| + (1 + \eta_0) \max_{1 \leq k \leq j-1} |v_k| \quad , \quad j > 1,$$

while $|v_1| \leq |v_0| + \Delta t|F_1|$. Since $(1 + \eta_0)^{-1} \leq 1$ (see Lemma 2.i), we can apply Lemma 3. Summing the terms and using Lemma 2.ii, we obtain

$$|v_i| \leq \left(1 + \sum_{j=2}^{i} \eta_{j-1}\right) |v_0| + \sum_{j=1}^{i} \Delta t|F_j| \leq C_\tau(t_i)|v_0| + \sum_{j=1}^{i} \Delta t|F_j|, \quad 1 \leq i \leq n ,$$

and, hence, the conclusion of the lemma. $\square$

Note that, if we set $F_j \doteq \hat{f}_j$ in (15), then Lemma 4 gives us directly the estimate

$$|u_i| \leq C_\tau(t_i)|u_0| + \sum_{j=1}^{i} \Delta t|\hat{f}_j| \leq C_\tau(t_i)|u_0| + \sum_{j=1}^{i} \int_{I_j} |f(s)|ds , \quad 1 \leq i \leq n ,$$

for the solution of (14). Hence, the following result holds.

Corollary . *If Δt is small, then the solutions of (14) satisfy*

$$|u_i| \leq C_\tau(t_i)|u_0| + \| f \|_{1,0,t_i} , \quad 1 \leq i \leq n .$$

3. CONVERGENCE RESULT

In this section we define the fully discrete scheme and prove our main result, i.e., that this scheme is convergent. The fully discrete scheme is based on the difference scheme (14), which we apply to the weak form of the problem (1)–(3) ,

$$(17) \qquad (u_t, v) + (Lu_t, v) + a(u, v) = (f, v) \quad , \quad \forall v \in V, \forall t \in I ,$$

$$(18) \qquad\qquad\qquad (u(0), v) = (u_0, v) \quad , \quad \forall v \in V .$$

In what follows, we shall assume that

$$(19) \quad u \in C(I; H) \cap L^\infty(I; V) \text{ is the unique solution of (17)–(18)}$$

$$\text{and additionally } u \in C^2(\tilde{Q}), u_{tt} \in L^1(I; H^2(\Omega) \cap V).$$

Now we specify the appropriate spatial discretization of the problem. Consider a finite element space $V^h \subset V$ of piecewise linear functions based on a quasi–uniform

triangulation of the domain Ω with parameter h. (We apply the standard construction described, e.g., in the monograph [7].) The space V^h is of dimension N_h and is spanned by basis functions $\{\psi_i(\cdot)\}_{i=1}^{N_h}$. The mass and the stiffness matrices are defined respectively by

$$\mathbf{C} \stackrel{\text{def}}{=} \{C_{i,m}\}_{i,m=1}^{N_h} \ , \quad C_{i,m} = (\psi_i, \psi_m) \ ,$$

$$\mathbf{A} \stackrel{\text{def}}{=} \{A_{i,m}\}_{i,m=1}^{N_h} \ , \quad A_{i,m} = a(\psi_i, \psi_m).$$

Note that $\mathbf{C}$ as a Gram matrix is always nonsingular and $\mathbf{A}$ is positive definite. The matrix $\mathbf{A}$ is always at least positive semidefinite, even if we replace (2) by non–Dirichlet boundary conditions.

We define the approximation $U : I \to V^h$ of u as a function which is piecewise linear in the time variable based on the set of values $U(t_j) \stackrel{\text{def}}{=} U_j \in V^h, 0 \le j \le n$. The latter are the solutions of the discrete problem

$$(20) \ (\partial_n U_j, v) + \sum_{k=1}^{j} \eta_{j-k}(\partial_n U_k, v) + a(U_j, v) \quad = \quad (\widehat{f_j}, v), \forall v \in V^h, 1 \le j \le n \ ,$$

$$(21) \qquad\qquad\qquad\qquad (U_0, v) \quad = \quad (u_0, v) \ , \forall v \in V^h \ .$$

The linear system (20) can be rewritten in the matrix–vector form

$$(22) \qquad \mathbf{B}\mathbf{U}_j = \Delta t \mathbf{F}_j + \mathbf{C}\mathbf{U}_{j-1} + \eta_{j-1}\mathbf{C}\mathbf{U}_0 + \sum_{k=1}^{j-1} \vartheta_{j-k}\mathbf{C}\mathbf{U}_k$$

where $\mathbf{B} \stackrel{\text{def}}{=} (1 + \eta_0)\mathbf{C} + \Delta t \mathbf{A}$, $\mathbf{F}^j \stackrel{\text{def}}{=} [F_1^j, F_2^j, \dots, F_{N_h}^j]^T$, $F_k^j = (\widehat{f_j}, \psi_k)$, and we identify $U_j \in V^h$ with $\mathbf{U}_j \in R^{N_h}$. With η_0 positive, the matrix $\mathbf{B}$ to be inverted is always nonsingular; hence, the approximations U_j, $0 \le j \le n$, are well defined.

Below we prove convergence of the approximation using the technique of elliptic projections introduced in [35] and apply the following well-known result, a consequence of the Nitsche argument (see, e.g., [7]).

Lemma 5. *Let $a(\cdot, \cdot)$ be defined as above. Then $a(\cdot, \cdot)$ is 0–regular and the projection $\tilde{u}$ of u is well defined by*

$$\tilde{u}(t) \in V^h : a(u(t) - \tilde{u}(t), v) = 0 \ , \quad \forall v \in V^h \ , \quad \forall t \in I.$$

If u satisfies (19), then $\tilde{u}$ is as smooth as u in the time variable, and the approximation properties

$$|u - \tilde{u}| \quad \le \quad \tilde{C}h^2 \parallel u \parallel_{H^2(\Omega)} \ , \quad \forall t \in \bar{I} \ ,$$

$$|u_t - \tilde{u}_t| \quad \le \quad \tilde{C}h^2 \parallel u_t \parallel_{H^2(\Omega)} \ , \quad \forall t \in \bar{I} \ ,$$

hold with a positive constant $\tilde{C}$ independent of h, u.

Now we prove the convergence rate estimate.

Theorem . *Let u satisfy (19). Let also (4), (5), (6) hold and Δt be sufficiently small. Then the approximation error satisfies the estimate*

$$(23) \qquad \sigma \stackrel{\text{def}}{=} \max_{1 \le j \le n} |u(t_j) - U_j| \le C(\Delta t + h^2)$$

with constant C independent of $h, \Delta t$.

Proof. The approximation error can be represented as a sum of two terms estimated individually by

$$\max_{1 \le j \le n} |u(t_j) - U_j| \le \max_{1 \le j \le n} |u(t_j) - \tilde{u}(t_j)| + \max_{1 \le j \le n} |\tilde{u}(t_j) - U_j| \le C_1 \left(\| u \|_{\infty,2,T} \right) h^2$$
$$+ C_2 \left(c_D, T, \tau, \| u_0 \|_{H^2(\Omega)}, \| u_t \|_{1,2,T}, \| u_{tt} \|_{1,0,T} \right) (\Delta t + h^2),$$

where we have used the approximation properties (see Lemma 5) to bound the first term and Lemma 6, to be proved below, for the second term. The constants C_1, C_2 depend on the problem data and its solution but are independent of the approximation parameters. Hence the inequality (23) follows. $\qquad\square$

We note that the convergence rate $O(\Delta t + h^2)$ predicted by the Theorem is of the same order as the corresponding rate for the analogous discrete scheme defined for the heat diffusion equation (i.e., if $\tau \equiv 0$). Such a result for the diffusion equation is known to be optimal (see [35]). Our assumptions on the smoothness of the solution (19) are only slightly stronger than those required for convergence of approximation to the diffusion equation.

Lemma 6. *Let the assumptions of the Theorem hold. Then each* $\xi_i = U_i - \tilde{u}(t_i)$, $1 \le i \le n$, *satisfies the estimate*

$$|\xi_i| \le \Delta t \left(c_D \| u_t \|_{1,2,t_i} + \mathcal{T}(t_i) \| u_{tt} \|_{1,0,t_i} \right)$$
$$+ h^2 C_\tau(t_i) \tilde{C} \left(\| u_0 \|_{H^2(\Omega)} + \| u_t \|_{1,2,t_i} \right) .$$

Proof. We omit the x-variable for simplicity of notation. Take $1 \le j \le n$ and $v \in V^h$ in (17), integrate both sides of the equation over I_j, use the definition of $\widehat{f}_j$ and subtract the result from (20) multiplied by Δt to get

$$(U_j - U_{j-1}, v) - (u(t_j) - u(t_{j-1}), v) + \sum_{k=1}^{j} \eta_{j-k}(U_k - U_{k-1}, v)$$
$$- \left(\int_{I_j} (Lu_t)(s) ds, v \right) + a(\Delta t U_j - \int_{I_j} u(s) ds, v) = 0 .$$

Using $U_k = \xi_k + \tilde{u}(t_k), 1 \le k \le j$, we rewrite the last identity as

$$(24) \quad (\xi_j - \xi_{j-1}, v) + \sum_{k=1}^{j} \eta_{j-k}(\xi_k - \xi_{k-1}, v) + \Delta t a(\xi_j, v)$$
$$= a(\alpha_j, v) + (u(t_j) - u(t_{j-1}), v) - (\tilde{u}(t_j) - \tilde{u}(t_{j-1}), v)$$
$$+ \left(\int_{I_j} (Lu_t)(s) ds, v \right) - \sum_{k=1}^{j} \eta_{j-k}(\tilde{u}(t_k) - \tilde{u}(t_{k-1}), v) ,$$

where we denote $\alpha_j = \int_{I_j} u(s) ds - \Delta t u(t_j)$. Note that from the definition of $\tilde{u}$ we have $(\mathbf{M}\alpha_j, v) = a(\alpha_j, v) = a(\int_{I_j} u(s) ds - \Delta t \tilde{u}(t_j), v)$. Now, in order to rewrite (24) in a more convenient form, we combine the last two terms on its right-hand side, using the fact that $\sum_{k=1}^{j} \eta_{j-k}(\tilde{u}(t_k) - \tilde{u}(t_{k-1})) = \int_{I_j} (L\bar{\tilde{u}}_t)(s) ds$. Here, $\bar{\tilde{u}}$ is the piecewise linear approximation (in the time variable) of $\tilde{u}$, and by $\bar{\tilde{u}}_t = (\bar{\tilde{u}})_t$ we mean its time derivative. (Note that $(\bar{\tilde{u}})_t = \widehat{(\tilde{u}_t)}$, i.e., $\bar{\tilde{u}}_t$ is piecewise constant.) We shall use the decomposition

$$u(s) - \tilde{\bar{u}}(s) = (u(s) - \bar{u}(s)) + (\bar{u}(s) - \tilde{\bar{u}}(s)) = g(s) + \tilde{g}(s),$$

from which, by $g(t_k) = u(t_k) - \bar{u}(t_k) = 0, 0 \leq k \leq n$, there follows

$$(u(t_j) - u(t_{j-1})) - (\tilde{u}(t_j) - \tilde{u}(t_{j-1})) = \int_{I_j} \tilde{g}_t(s) ds.$$

The identity (24) thus reads

$$(\xi_j - \xi_{j-1}, v) + \sum_{k=1}^{j} \eta_{j-k}(\xi_k - \xi_{k-1}, v) + \Delta t a(\xi_j, v)$$

$$= a(\alpha_j, v) + \left(\int_{I_j} \tilde{g}_t(s) ds, v \right) + \left(\int_{I_j} (Lg_t)(s) ds, v \right) + \left(\int_{I_j} (L\tilde{g}_t)(s) ds, v \right).$$

If we divide both sides of this identity by Δt, it takes the (weak) form of the discrete equation for ξ_j (compare e.g. (15), (16)) considered in Lemma 4, with the vector F_j defined by

$$(25) \qquad \Delta t F_j = \mathbf{M}\alpha_j + \int_{I_j} \left(\tilde{g}_t(s) + (Lg_t)(s) + (L\tilde{g}_t)(s) \right) ds.$$

Directly applying the result of Lemma 4, we get

$$(26) \qquad |\xi_i| \leq C_\tau(t_i)|\xi_0| + \sum_{j=1}^{i} \Delta t |F_j|,$$

and now we need to find bounds for ξ_0 and $\sum_{j=1}^{i} \Delta t |F_j|$. This is pursued below.

By definition of U_0 and the approximation property (Lemma 5), we have

$$|\xi_0| = |u_0 - \tilde{u}_0| \leq \tilde{C}h^2 \| u_0 \|_{H^2(\Omega)}.$$

For the first term in (25) we have

$$|\mathbf{M}\alpha_j| = |\int_{I_j} \mathbf{M}(u(s) - u(t_j)) ds| \leq \int_{I_j} |\mathbf{M}(u(s) - u(t_j))| ds$$

$$\leq c_D \int_{I_j} \| u(s) - u(t_j) \|_{H^2(\Omega)} ds \leq c_D \int_{I_j} \int_{t_j}^{s} \| u_t(r) \|_{H^2(\Omega)} dr ds$$

$$\leq c_D \Delta t \int_{I_j} \| u_t(s) \|_{H^2(\Omega)} ds.$$

For the second term in (25) we apply Lemma 5 to get

$$\left| \int_{I_j} \tilde{g}_t(s) ds \right| = \left| \int_{I_j} \left(\frac{u(t_j) - u(t_{j-1})}{\Delta t} - \frac{\tilde{u}(t_j) - \tilde{u}(t_{j-1})}{\Delta t} \right) ds \right|$$

$$= \left| \int_{I_j} (u_t(s) - \tilde{u}_t(s)) ds \right| \leq \int_{I_j} |u_t(s) - \tilde{u}_t(s)| ds \leq \tilde{C}h^2 \int_{I_j} \| u_t(s) \|_{H^2(\Omega)} ds.$$

For the remaining terms we have

$$\left| \int_{I_j} (Lg_t)(s)ds \right| \le \int_{I_j} (L|g_t|)(s)ds,$$

$$\left| \int_{I_j} (L\tilde{g}_t)(s)ds \right| \le \int_{I_j} (L|\tilde{g}_t|)(s)ds.$$

The desired bounds follow from continuity of L (its norm on $L^1((0, t_i))$ equals $\mathcal{T}(t_i)$) combined with estimates for $|g_t|$ and $|\tilde{g}_t|$. These are the error between u_t and its piecewise constant approximant (in the time variable) $\widehat{(u_t)} = (\bar{u})_t$, and the error between $(\bar{u})_t$ and its elliptic projection $\widehat{(\tilde{u})}_t = (\bar{\tilde{u}})_t$, respectively. The estimates for the latter then follow from Lemma 5, while the former is dealt with in a standard manner:

$$\sum_{j=1}^{i} \int_{I_j} (L|g_t|)(s)ds = \int_0^{t_i} (L|g_t|)(s)ds \le \mathcal{T}(t_i) \int_0^{t_i} |g_t(s)|ds \le \mathcal{T}(t_i)\Delta t \parallel u_{tt} \parallel_{1,0,t_i},$$

$$\sum_{j=1}^{i} \int_{I_j} (L|\tilde{g}_t|)(s)ds = \int_0^{t_i} (L|\tilde{g}_t|)(s)ds \le \mathcal{T}(t_i) \sum_{j=1}^{i} \int_{I_j} |\tilde{g}_t(s)|ds$$

$$\le \mathcal{T}(t_i)\tilde{C}h^2 \sum_{j=1}^{i} \int_{I_j} \parallel \bar{u}_t(s) \parallel_{H^2(\Omega)} ds$$

$$= \mathcal{T}(t_i)\tilde{C}h^2 \sum_{j=1}^{i} \int_{I_j} \parallel \frac{u(t_j) - u(t_{j-1})}{\Delta t} \parallel_{H^2(\Omega)} ds$$

$$\le \mathcal{T}(t_i)\tilde{C}h^2 \sum_{j=1}^{i} \parallel \int_{I_j} u_t(s)ds \parallel_{H^2(\Omega)} \le \mathcal{T}(t_i)\tilde{C}h^2 \int_0^{t_i} \parallel u_t(s) \parallel_{H^2(\Omega)} ds.$$

Hence, (26) leads to the estimate

$$|\xi_i| \le C_\tau(t_i)\tilde{C}h^2 \parallel u_0 \parallel_{H^2(\Omega)} + c_D\Delta t \parallel u_t \parallel_{1,2,t_i} + \tilde{C}h^2 \parallel u_t \parallel_{1,2,t_i}$$
$$+ \Delta t \mathcal{T}(t_i) \parallel u_{tt} \parallel_{1,0,t_i} + \mathcal{T}(t_i)\tilde{C}h^2 \parallel u_t \parallel_{1,2,t_i},$$

which completes the proof of the lemma. $\qquad\square$

4. NUMERICAL EXPERIMENTS

In order to verify the convergence rate estimates derived above, we performed numerical experiments in which the approximation error was measured for a given exact solution and an approximate solution calculated with the help of the discrete scheme (22). Our tests were made for two differential problems: the linear one as in the original problem (1)–(3) and a quasilinear one (see Example 3 below). In both cases we obtained a confirmation of the convergence rate predicted by the Theorem for the linear case. The convergence for the quasilinear problem is important in applications, e.g., it was used in calculations realized for coupled nonlinear models of flow in fissured media (see [26, 27]).

We choose $\Omega = (0, 1) \times (0, 1), T = 1$, and consider a function

$$u(x, t) = \sin(\pi x_1) \sin(\pi x_2) u_f(t), \quad x = (x_1, x_2) \in \Omega, t \in I,$$

which is the exact solution of the problem (1)–(3) with $u_f(\cdot)$ defined below. Note that u satisfies the homogeneous Dirichlet boundary condition on Γ and that u_0 is

TABLE 1

Examples 1 and 2				Example 3			
i	h_i	Δt_i	n_i	i	h_i	Δt_i	n_i
1	1/11	0.02	50	1	1/11	0.02	50
2	1/15	0.0102041	98	2	1/15	0.0102041	98
3	1/20	0.00552486	181	3	1/17	0.0078125	128
4	1/30	0.0023753	421	4	1/20	0.00552486	181
5	1/50	0.000832639	1201	5	1/30	0.0023753	421

TABLE 2

i	τ_1		τ_2		τ_3	
	σ_i	ς_i	σ_i	ς_i	σ_i	ς_i
1	0.0121171	1.9807	0.0114434	1.98187	0.00389853	2.00996
2	0.00625372	1.9935	0.00590208	1.99396	0.00197916	2.00572
3	0.00341042	2.02231	0.0032175	2.02016	0.00106884	1.9633
4	0.00147038	1.98141	0.00138694	1.98251	0.000458784	2.00928
5	0.000516142	1.99029	0.000486806	1.99084	0.000160696	2.00374

determined by the choice of u_f. The convolution kernel τ is one of the following three arising in applications (see [14, 25]).

(a) $\tau_1(t) \equiv 0$,
(b) $\tau_2(t) = e^{-t}$,
(c) $\tau_3(t) = 6\sum_{k=1}^{\infty} e^{-k^2\pi^2 t}$.

Note that in case (a) the problem (1)–(3) is equivalent to the heat equation, and in case (b) we are given a nonsingular kernel. In case (c), τ_3 is the indicated Fourier–Bessel series which is singular at the origin, and its values cannot be calculated exactly but only approximately. This is done by deleting the terms in the tail of the series, i.e., the summation includes only the first N_τ terms. Such an approximation gives satisfactory results when calculating values of τ or of its integral $\mathcal{T}$. However, it is not sufficient if we need values of any convolution integral $\tau * g$ unless g is constant.

The spatial domain Ω is covered with a regular triangular finite element grid with discretization parameter h taking test values $h_i, 1 \le i \le 5$ as in Table 1 (different test values for different examples). For each i we set the time step $\Delta t_i \approx h_i^2$ and set the appropriate number of time steps n_i; next we solve (22) and calculate the approximation error σ_i defined by (23). In order to study the convergence rate, we first find $\tilde{h}$ and $\tilde{\sigma}$, which are geometric means of the values of h_i, σ_i, respectively, and then we identify the exponent ς_i according to the conjecture

$$(\frac{h_i}{\tilde{h}})^{\varsigma_i} = \frac{\sigma_i}{\tilde{\sigma}}.$$

Note that the Theorem predicts $\varsigma_i = 2$.

Example 1. The function u_f is set to $u_f(t) = t+1$ and $\mathbf{D} \equiv \mathbf{I}$, the identity tensor. The function f in (1)–(3) can be calculated analytically for each of the kernels τ_1, τ_2, τ_3. The corresponding approximate error and values of σ_i, ς_i are presented in Table 2.

TABLE 3

function u_f	i	τ_1		τ_2	
		σ_i	ς_i	σ_i	ς_i
$f_u(t) = e^t$	1	0.000706362	1.90566	0.000697771	1.9044
	2	0.000380489	1.96338	0.000376127	1.96284
	3	0.000211804	2.12632	0.000209504	2.13106
	4	9.27716e-05	1.9087	9.17606e-05	1.90763
	5	3.2836e-05	1.95219	3.24782e-05	1.95177
$f_u(t) = e^{-t}$	1	0.0213288	1.98014	0.0200613	1.9818
	2	0.0110132	1.99367	0.0103489	1.99434
	3	0.00600413	2.01945	0.00563905	2.01641
	4	0.00258976	1.98084	0.00243161	1.98243
	5	0.000909263	1.98983	0.000853619	1.99061

TABLE 4

i	τ_1		τ_2		τ_3	
	σ_i	ς_i	σ_i	ς_i	σ_i	ς_i
1	0.0065474	1.99746	0.0065474	2.05032	0.00369569	1.93658
2	0.00333973	1.99079	0.00323035	1.95026	0.00193864	1.97514
3	0.00255872	1.9744	0.00247508	1.76986	0.00149591	2.11339
4	0.00181381	2.00883	0.00175457	2.05651	0.00106676	1.93342
5	0.000783341	1.99133	0.000757843	2.00294	0.000464253	1.95902

The results confirm the convergence claims of the Theorem. The approximation error σ_i decays with $h_i^{\varsigma_i}$ and ς_i is close to 2, i.e., the rate of convergence is quadratic. Note that there is no essential difference between the results corresponding to τ_1 (i.e., the case of no memory effects) and the results corresponding to τ_2 or τ_3. Also, the approximation scheme gives results of the same quality for both kernels τ_2 and τ_3, i.e., the singularity of the kernel does not slow the convergence of the scheme.

Example 2. With $\mathbf{D}$ as in Example 1, the function u_f is set to be $u_f(t) = e^t$ or $u_f(t) = e^{-t}$; the corresponding function f in (1)–(3) can be calculated analytically only for nonsingular kernels, i.e., for τ_1 and τ_2. The results collected in Table 3 confirm the convergence claims.

Example 3. In this example the differential problem to be solved is the quasilinear equation

$$u_t + Lu_t - \nabla \cdot (\mathbf{D}(u)\nabla u) = f(x,t).$$

Elements of the tensor $\mathbf{D}$ are dependent on the solution u by $D_{i,j}(u) = \delta_{i,j}(u + 1)$ with Kronecker's symbol $\delta_{i,j}$. The corresponding numerical approximation is defined in a way analogous to the linear case. The only modification concerns resolution of the dependence of the coefficient stiffness matrix $\mathbf{A}$ in (22) on the solution. This is achieved by the standard predictor–corrector procedure (see [35]) with two predictor–corrector steps. The approximation procedure is convergent, as is shown by the values of ς_i calculated for $u_f(t) = 1 + t$ as in Table 4 and for $u_f(t) = e^t, u_f(t) = e^{-t}$ as in Table 5.

TABLE 5

function u_f	i	τ_1		τ_2	
		σ_i	ς_i	σ_i	ς_i
	1	0.00278412	1.96864	0.00270949	1.97002
	2	0.00144114	1.99197	0.00140062	1.98802
$f_u(t) = e^t$	3	0.0011072	2.06576	0.00107681	2.06676
	4	0.000785405	1.9825	0.000763785	1.98285
	5	0.00034055	1.97773	0.000331185	1.97777
	1	0.0141327	2.04897	0.0137326	2.04952
	2	0.00704049	2.00435	0.00683974	2.00475
$f_u(t) = e^{-t}$	3	0.00536171	1.85908	0.00520845	1.85889
	4	0.00377781	2.07936	0.00366939	2.08025
	5	0.00162383	2.01724	0.00157709	2.01761

5. IMPLEMENTATION

There are two main complexity issues in the implementation of the numerical algorithm for solution of the problem (22). The first is the amount of computer memory necessary to keep track of the "history" of the process. The second is related to the time spent on the solution of the linear system (22) which can be decreased via an implementation in a multiprocessing environment. Below we discuss both issues.

The history of the process expressed by the convolution integral in the continuous problem requires that the values of $u(t)$ for $0 < t < T$ depend on the values of the derivative of the solution $u_t(s)$ in the whole interval $s \in (0, t)$. In the language of the discrete solution it means that in order to calculate the right-hand side of (22) at the jth time layer, all the vectors $\mathbf{U}_k, 0 \le k < j$ have to be "accessible" to the numerical algorithm, so we must keep them stored in memory. This requirement dramatically increases the amount of memory necessary for the implementation of the algorithm. In contrast, in the case of the heat equation one needs only to store values from the previous time layer, i.e., only one solution vector instead of j vectors as in the "memory effects" case.

The storage problem, as an important complexity issue, had been seriously considered in [20, 24, 31, 34]. These theoretical results apply to the integrals with bounded kernels, except for those presented in [34], which however rely on an explicit formula for the kernel. As mentioned in the Introduction, the convolution kernels for the class of problems considered in this paper arise from some auxiliary differential problem and are usually singular. Their values in general are available only by numerical approximation, see [25, 27]). The general solution for the storage issue applicable in the singular kernel case is still to be sought. Below we supply some remarks concerning the practical solution to this issue.

One way to avoid storing all the time layer vectors would be to explicitly cut off in (22) all the calculations involving time layers which are "far enough" from the current time layer. Practically, we do this by setting the coefficients $\vartheta_r = 0$ either if (a) $\vartheta_r < \epsilon$ or if (b) $r > n_\epsilon$; the barrier parameters (n_ϵ, ϵ) are set appropriately to the data of the problem. In other words, we use perturbed values of coefficients ϑ_r. The well-posedness of the problem implies that the perturbed values of τ

introduce a continuous perturbation in the approximation error. However, the values of ϑ_r for fixed Δt are related rather to the derivative of τ and not to τ, and our numerical experiments show that one should be very careful with deciding which time layers are to be taken into account or ignored while calculating the "history of the diffusion." If we follow approach (a), then in the extreme cases a careless choice of ϵ may even prevent convergence. With decreasing discretization parameters the barrier value for ϑ_r is very soon reached, even for a singular kernel. Then the number of layers taken into account may not be big enough to guarantee convergence. Hence, an individual choice of appropriate ϵ fitting the discretization parameters and convolution kernel is necessary.

In the case (b) the best strategy, according to our numerical experiments, is to set a fixed ratio between the number of layers to be included in calculations and the total number of layers, for example $1/2$. Then the approximation error decays, but the price paid for the memory savings is slower convergence than that predicted by the Theorem.

Another approach to this issue might be to use some implementation-specific tricks that avoid the direct storage in the computer memory. In certain computing environments some cache or register memory devices (faster than the disk storage devices) are available, which might provide secondary storage for the time layer vectors. Additionally, in a multiprocessing environment, a process calculating the discrete convolution term using the time layer vectors stored aside can execute in parallel to other components of the algorithm.

The second complexity issue, i.e., that of the linear system solver, can be efficiently resolved with use of a multiprocessing computing environment as well. We describe our results concerning this issue below.

The linear system (22) is a system with a sparse positive definite matrix **B**. We solve it by the Schwarz Alternative Method of Additive Type (SAM) which is a domain decomposition (DD) method belonging to parallelizable PCG (preconditioned conjugate gradient) methods (see [8, 9, 10] and references therein). In this method the spatial domain of the differential problem, which is discretized with a fine grid, is split into subdomains forming a coarse grid. The problem is delegated to solvers executing in parallel and corresponding to the overlapping subdomains, with an additional solver operating on the coarse mesh and providing the global exchange of information between subdomains. The method is effective for linear systems arising from elliptic 2D and 3D problems (see [11] or [5]) as well as for parabolic problems, where in most cases no coarse mesh solver is necessary (see [4, 8, 9]). The algorithm can be implemented very well in multiprocessing environments (see [11, 13]).

Our results show that SAM applies equally well to the problem (22) as it does to the algebraic problems arising from the heat equation. This is to be expected since the performance of the method depends on the conditioning of the linear system matrix **B**, which is even better for nontrivial τ than for $\tau \equiv 0$. For details on implementation, see [25].

6. APPENDIX

Below we prove the continuous stability result for the general case of possibly singular convolution kernels. This result is given for completeness, since in Section 2 we considered (for the sake of exposition) only the case of kernels defined at the

origin. The proof relies on a particular property of the convolution operator, and this is demonstrated in the sequel.

Lemma 7. *Let τ satisfy (4). Let also u be a solution of (7)–(8). Then u satisfies*

$$|u(t)| \leq C_\tau(t)|u_0| + C \parallel f \parallel_{2,0,t} .$$

Proof. Multiply (7) by u and integrate the resulting identity over Ω to obtain

$$\frac{1}{2}\frac{d}{dt}|u|^2 + (Lu_t, u) + (Mu, u) = (f, u) \quad , \quad t > 0.$$

This identity, after integration over $(0,t)$, gives

$$(27) \qquad \frac{1}{2}|u(t)|^2 - \frac{1}{2}|u_0|^2 + \mathcal{Q}_H(t,\tau,u) + \int_0^t (\mathbf{M}u, u)ds = \int_0^t (f, u)ds,$$

where we have denoted

$$(28) \qquad \mathcal{Q}_H(t,\tau,u) \overset{\text{def}}{=} \int_0^t (Lu_t, u)_H ds = \int_0^t ((\tau * u_t)(s), u(s))ds.$$

Lemma 8 contains the result giving us the lower bound $\mathcal{Q}_H(t,\tau,u) \geq -\frac{\mathcal{T}(t)}{2}|u(0)|^2$. On the other hand, V–coercivity of the form $a(\cdot,\cdot)$ associated with $\mathbf{M}$ yields $(\mathbf{M}u, u) = a(u,u) \geq c_a \parallel u \parallel^2 \geq \frac{c_a}{c_P^2}|u|^2$. Combining the estimates above and using the Schwarz inequality, we get

$$\frac{1}{2}|u(t)|^2 + \frac{c_a}{c_P^2}\int_0^t |u(s)|^2 ds \leq \frac{1}{2}|u_0|^2(\mathcal{T}(t)+1) + \frac{1}{2\epsilon}\int_0^t |f(s)|^2 ds + \frac{\epsilon}{2}\int_0^T |u(s)|^2 ds.$$

With $\epsilon = 2\frac{c_a}{c_P^2}$ in the inequality $ab \leq \frac{1}{2\epsilon}a^2 + \frac{\epsilon}{2}b^2$, we conclude

$$|u(t)|^2 \leq C_\tau(t)|u_0|^2 + \frac{c_P^2}{2c_a}\int_0^t |f(s)|^2 ds,$$

hence, the conclusion of the lemma.   □

In the lemma below we adopt notation for functions and measures as in [12], where general properties of monotone kernels are considered. In particular, we use v' to denote the derivative of a (locally absolutely continuous) function v with respect to the time variable. The proof of the lemma follows the idea sketched in [12]. (See for example Lemma 18.4.1, p. 574, in this reference.)

Lemma 8. *Let τ satisfy (4). Then for $v \in H^1(I; H)$ and $\mathcal{Q}_H(t,\tau,v)$ defined by (28) it follows that*

$$(29) \qquad \mathcal{Q}_H(t,\tau,v) \geq -\frac{\mathcal{T}(t)}{2}|v(0)|^2.$$

Proof. First we prove the counterpart of (29) for real-valued functions $v \in H^1(I)$ and $\mathcal{Q}_\mathbf{R}(t,\tau,v)$, i.e., we want to prove

$$(30) \qquad \mathcal{Q}_{\mathbf{R}}(t, \tau, v) = \int_0^t (\tau * v')(s)v(s)ds \geq -\frac{\mathcal{T}(t)}{2}(v(0))^2.$$

Note that, if τ is nonsingular, we can write $\tau * v' = -\tau(t)v(0) + \tau(0)v(t) + \tau' * v$. In the general case of τ singular at the origin we reformulate $\tau * v'$ similarly, using the (Borel) measure γ obtained as the (distribution) derivative of τ. Formally, we have $\gamma((0,t)) = \tau(t)$ at all points of continuity of τ, i.e., for $t > 0$. This implies $(f * \tau)' = f * \gamma$ for any $L^2(I)$ integrable f, and in addition we derive

$$\tau * v' = -\tau(t)v(0) + (v * \tau)' = -\tau(t)v(0) + v * \gamma.$$

Now rewrite $\mathcal{Q}_{\mathbf{R}}(t, \tau, v)$ and apply the identity $ab = \frac{a^2}{2} + \frac{b^2}{2} - \frac{1}{2}(a - b)^2$ to get

$$(31) \quad \mathcal{Q}_{\mathbf{R}}(t, \tau, v) = \int_0^t (v' * \tau)(s)v(s)ds$$

$$= -\int_0^t v(0)v(s)\tau(s)ds + \int_0^t \int_{[0,s]} v(s)v(s - r)\gamma(dr)ds$$

$$= -\frac{1}{2}\int_0^t v^2(0)\tau(s)ds - \frac{1}{2}\int_0^t v^2(s)\tau(s)ds + \frac{1}{2}\int_0^t (v(0) - v(s))^2\tau(s)ds$$

$$+ \frac{1}{2}\int_0^t v^2(s)\int_{[0,s]} \gamma(dr)ds + \frac{1}{2}\int_0^t \int_{[0,s]} v^2(s - r)\gamma(dr)ds$$

$$- \frac{1}{2}\int_0^t \int_{[0,s]} (v(s) - v(s - r))^2\gamma(dr)ds.$$

The definition of γ implies that the second and the fourth terms cancel. Moreover, for the fifth term we have

$$\int_0^t \int_{[0,s]} v^2(s - r)\gamma(dr)ds = \int_0^t (v^2 * \gamma)(s)ds$$

$$= \int_0^t (v^2 * \tau)'(s)ds = \int_0^t v^2(t - s)\tau(s)ds,$$

so we can write

$$(32) \quad \mathcal{Q}_{\mathbf{R}}(t, \tau, v) = -\frac{1}{2}v^2(0)\int_0^t \tau(s)ds + \frac{1}{2}\int_0^t (v(0) - v(s))^2\tau(s)ds$$

$$+ \frac{1}{2}\int_0^t v^2(t - s)\tau(s)ds - \frac{1}{2}\int_{[0,t]} \int_s^t (v(r) - v(r - s))^2 dr\gamma(ds),$$

where the last term is obtained from the sixth term in (31) by a change of the order of integration.

With the first term on the right-hand side of (32) equal to our desired lower bound in (30), it remains to prove that the remaining terms on the right-hand side of (32) are nonnegative. This observation is trivial for the second and the third terms. For the last term it follows from the fact that on $(0,t)$ the measure γ is negative. Also, although γ contains a positive point mass at zero, at $s = 0$ the value of the integrand $(v(r) - v(r - s))$ becomes 0. Hence (30) is proved.

In order to obtain (29), we integrate the inequality (30) (it is satisfied for $v(x)$ at every $x \in \Omega$ treated as a parameter) over Ω and get the desired inequality involving the scalar product in H. $\qquad\square$

REFERENCES

1. T. Arbogast, *Analysis of the Simulation of Single Phase Flow Through a Naturally Fractured Reservoir*, SIAM J. Numer. Anal. **26** (1989), 12-29. MR **90e**:76122
2. T. Arbogast, J. Douglas (Jr.), *Dual–Porosity Models for Flow in Naturally Fractured Reservoirs*, In "Dynamics of Fluids in Hierarchical Porous Media," J. H. Cushman, ed., Academic Press, London, 1990, 177–221.
3. T. Arbogast, J. Douglas (Jr.), U. Hornung, *Derivation of the Double Porosity Model of Single Phase Flow via Homogenization Theory*, SIAM J. Math. Anal. **21** (1990), 823-836. MR **91d**:76074
4. Xiao Chuan Cai, *Additive Schwarz algorithms for parabolic convection–diffusion equations*, Numer. Math. **60** (1991), 41-61. MR **93a**:65127
5. Xiao Chuan Cai, W.D. Gropp, D.E. Keyes, *Convergence rate estimate for a domain decomposition method*, Numer. Math. **61** (1992), 153–169. MR **92k**:65181
6. C. Chen, V. Thomée, L. B. Wahlbin, *Finite element approximation of a parabolic integro-differential equation with a weakly singular kernel*, Math. Comp. **58** (1992), 587-602. MR **93g**:65120
7. P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, (1978) North–Holland. MR **58**:25001
8. M. Dryja, O. Widlund, *An Additive Variant of the Schwarz Alternating Method for the Case of Many Subregions*, Technical Report 339, Department of Computer Science, Courant Institute of Mathematical Sciences, New York, December 1987.
9. M. Dryja, O.B. Widlund, *Some Domain Decomposition Algorithms for Elliptic Problems*, in: Iterative Methods for Large Linear Systems, Academic Press, 1990. CMP 90:07
10. Domain Decomposition Methods for Partial Differential Equations, Proceedings of the First International Symposium on Domain Decomposition Methods for Partial Differential Equations, *R. Glowinski, G. H. Golub, G. A. Meurant, J. Périaux eds.*, Paris, France, January, 1987, SIAM, Philadelphia, 1988.
11. A. Greenbaum, Congming Li, Han Zheng Cao, *Parallelizing Preconditioned Conjugate Gradient Algorithms*, Technical Report, Courant Institute, 1988.
12. G. Gripenberg, S.-O. Londen, O. Staffans, *Volterra Integral and Functional Equations*, Cambridge University Press, Cambridge, 1990. MR **91c**:45003
13. K.H. Hoffmann, J. Zou, *Parallel efficiency of domain decomposition methods*, Parallel Computing **19** (1993), 1375-1392.
14. U. Hornung, R. E. Showalter, *Diffusion Models for Fractured Media*, Jour. Math. Anal. Appl. **147** (1990), 69-80. MR **91d**:76072
15. Y. Lin, V. Thomée, L. B. Wahlbin, *Ritz-Volterra projections to finite element spaces and applications to integrodifferential and related eqautions*, SIAM J. Numer. Anal. **28** (1991), 1047-1070. MR **92k**:65193
16. P. Linz, *Analytical and Numerical Methods for Volterra Equations*, (1985) SIAM, Philadelphia. MR **86m**:65163
17. C. Lubich, *Convolution Quadrature and Discretized Operational Calculus, Parts I & 2*, Numer. Math. **52** (1988), 129-145 & 413-425. MR **89g**:65018; MR **89g**:65019

18. R. C. MacCamy, J. S. Wong, *Stability Theorems for Some Functional Equations*, Trans. Amer. Math. Soc. **164** (1972), 1-37. MR **45**:2432
19. Maria-Luisa Mascarenhas, *A linear homogenization problem with time dependent coefficient*, Trans. Amer. Math. Soc **281** (1984), 179-195. MR **85c**:45002
20. V. McLean, V. Thomée, L.B. Wahlbin, *Discretization with variable time steps of an evolution equation with a positive type memory term*, Applied Mathematics Report AMRR 93.18, December 1993 School of Math., The University of New South Wales.
21. R. K. Miller, *An integrodifferential equation for rigid heat conductors with memory*, Jour. Math. Anal. Appl. **66** (1978), 313-332. MR **80g**:45015
22. B. Neta, *Numerical Solution of a Nonlinear Integro–differential Equation*, Jour. Math. Anal. Appl. **89** (1982), 598-611. MR **84a**:65105
23. J. W. Nunziato, *On heat conduction in materials with memory*, Quarterly Appl. Math. **29** (1971), 187–204. MR **45**:4749
24. A.K. Pani, V. Thomée, L.B. Wahlbin, *Numerical methods for hyperbolic and parabolic integrodifferential equations*, J. Integral Equations Appl. **4** (1992), 533–584. MR **94c**:65167
25. M. Peszyńska, *Fluid Flow Through Fissured Media. Mathematical Analysis and Numerical Approach*, Ph. D. Thesis (1992), University of Augsburg.
26. M. Peszyńska, *Finite element approximation of a model of nonisothermal flow through fissured media*, in: Finite Element Methods, M. Křižek, P. Neittaanmäki, R. Stenberg (Eds), Marcel Dekker, 1994, 357–366.
27. M. Peszyńska, *On a model for nonisothermal flow in fissured media*, Differential Integral Equations **8** (1995), 1497–1516. CMP 95:12
28. M. Peszyńska, *Analysis of an integro–differential equation arising from modelling of flows with fading memory through fissured media*, J. Partial Diff. Eqs. **8** (1995), 159–173. MR **96a**:45007
29. A.H. Schatz, V. Thomée, W.L. Wendland, *Mathematical Theory of Finite and Boundary Element Methods*, Birkhäuser, Basel–Boston–Berlin, 1990. MR **92f**:65004
30. R. E. Showalter, *Distributed Microstructure Models of Porous Media*, in: "Flow in Porous Media: proceedings of the Oberwolfach conference, June 21–27, 1992", J. Douglas Jr. and U. Hornung, eds., Birkhäuser, Basel, 1993, 155–163. MR **95a**:76091
31. I. H. Sloan, V. Thomée, *Time Discretization of an Integro–Differential Equation of Parabolic Type*, SIAM J. Numer. Anal. **23** (1986) 1052–1061. MR **87j**:65113
32. L. Tartar, *Nonlocal effects induced by homogenization*, in: "Partial Differential Equations and the Calculus of Variations, Essays in Honor of Ennio de Giorgi," F. Colombini et. al., eds., Birkhäuser, Boston, 1989, 925-938. MR **91c**:35018
33. L. Tartar, *Memory effects and homogenization*, Arch. Rat. Mech. Anal. **111** (1990), 121–133. MR **92h**:35019
34. V. Thomée, L. B. Wahlbin, *Long time numerical solution of a parabolic equation with memory*, Dept. of Math, Chalmers University of Technology, The University of Göteborg, Preprint No 1992-12/ISSN 0347-2809
35. M. F. Wheeler, *A Priori L_2 Error Estimates for Galerkin Approximations to Parabolic Partial Differential Equations*, SIAM J. Numer. Anal., **10** (1973), 723-759. MR **50**:3613
36. Nai-ying Zhang, *On fully discrete Galerkin approximations for partial integro–differential equations of parabolic type*, Math. of Comp. **60** (1993) 133-166. MR **93d**:65088

SYSTEMS RESEARCH INSTITUTE, POLISH ACADEMY OF SCIENCES, UL. NEWELSKA 6   01-447 WARSZAWA, POLAND
E-mail address: `mpesz@ibspan.waw.pl`