

ERROR ANALYSIS OF A FINITE DIFFERENCE METHOD ON GRADED MESHES FOR A TIME-FRACTIONAL DIFFUSION EQUATION*

MARTIN STYNES[†], EUGENE O’RIORDAN[‡], AND JOSÉ LUIS GRACIA[§]

Abstract. A reaction-diffusion problem with a Caputo time derivative of order $\alpha \in (0, 1)$ is considered. The solution of such a problem is shown in general to have a weak singularity near the initial time $t = 0$, and sharp pointwise bounds on certain derivatives of this solution are derived. A new analysis of a standard finite difference method for the problem is given, taking into account this initial singularity. This analysis encompasses both uniform meshes and meshes that are graded in time, and includes new stability and consistency bounds. The final convergence result shows clearly how the regularity of the solution and the grading of the mesh affect the order of convergence of the difference scheme, so one can choose an optimal mesh grading. Numerical results are presented that confirm the sharpness of the error analysis.

Key words. fractional differential equation, initial boundary value problem, weak singularity, L1 scheme, graded mesh

AMS subject classifications. 65M06, 35B65

DOI. 10.1137/16M1082329

1. Introduction. The class of time-dependent problems with a fractional temporal derivative of order $\alpha \in (0, 1)$ has attracted much attention in recent years, since these problems arise in various models of physical processes; see the references in [8, 12, 17]. In the present paper we shall present a new and incisive analysis of a finite difference method for solving the most widely studied problem in this class.

Consider the initial boundary value problem

$$(1.1a) \quad D_t^\alpha u - p \frac{\partial^2 u}{\partial x^2} + c(x)u = f(x, t)$$

for $(x, t) \in Q := (0, l) \times (0, T]$ with

$$(1.1b) \quad u(0, t) = 0, \quad u(l, t) = 0 \quad \text{for } t \in (0, T],$$

$$(1.1c) \quad u(x, 0) = \phi(x) \quad \text{for } x \in [0, l],$$

where $0 < \alpha < 1$, p is a positive constant, $c \in C[0, l]$ with $c \geq 0$, $f \in C(\bar{Q})$, where $\bar{Q} := [0, l] \times [0, T]$, and $\phi \in C[0, l]$. More regularity hypotheses will be imposed on

*Received by the editors June 30, 2016; accepted for publication (in revised form) January 11, 2017; published electronically April 25, 2017.

<http://www.siam.org/journals/sinum/55-2/M108232.html>

Funding: The research of the first author was supported in part by the National Natural Science Foundation of China under grants 91430216 and NSAF-U1530401. The research of the third author was partly supported by the Institute of Mathematics and Applications (IUMA), the project MTM2016-75139-R, and the Diputación General de Aragón.

[†]Applied and Computational Mathematics Division, Beijing Computational Science Research Center, Beijing 100094, China (m.stynes@csrc.ac.cn).

[‡]School of Mathematical Sciences, Dublin City University, Glasnevin, Dublin 9, Ireland (eugene.oriordan@dcu.ie).

[§]Department of Applied Mathematics, Torres Quevedo Building, Campus Rio Ebro, University of Zaragoza, 50018 Zaragoza, Spain (jlgracia@unizar.es).

these functions in section 2. Furthermore, in (1.1a) D_t^α denotes the *Caputo fractional derivative* which is defined [5] by

$$D_t^\alpha g(x, t) := \frac{1}{\Gamma(1-\alpha)} \int_{s=0}^t (t-s)^{-\alpha} \frac{\partial g(x, s)}{\partial s} ds \quad \text{for } (x, t) \in Q.$$

Luchko [16] proves existence and uniqueness of a classical solution u to (1.1), i.e., $D_t^\alpha u$ and $p\partial^2 u/\partial x^2$ both exist pointwise inside Q and u satisfies (1.1) pointwise, under additional regularity and corner compatibility hypotheses that we shall discuss in section 2. (For a discussion of existence and uniqueness of the solution of (1.1) in a Sobolev space setting, see [12].)

More general boundary conditions $u(0, t) = \phi_L(t)$ and $u(l, t) = \phi_R(t)$, where ϕ_L and ϕ_R are smooth, are easily reduced to the homogeneous case of (1.1b) by a standard linear change of variable.

It is important to note that the smoothness of all the data of (1.1) does *not* imply smoothness of the solution u in the closed domain \bar{Q} ; this phenomenon will be discussed in more detail in section 2. The essential feature of all typical solutions of (1.1) is that u has an initial layer (see Figure 1) at $t = 0$ and $\partial u(x, t)/\partial t$ blows up as $t \rightarrow 0^+$.

Many papers have considered finite difference methods for the solution of (1.1) on uniform meshes; see the references in [2, 10]. But unlike [2, 10], most papers discussing (1.1) have ignored the possible presence of an initial layer in the solution at $t = 0$ and have presented convergence analyses that make the unrealistic assumption that u is smooth on the closed domain \bar{Q} . In [10, sections 1.2 and 2.3] the effect of this layer on the convergence rates claimed in the literature is examined in detail; as the authors point out, smoothness assumptions made in the literature are “not always valid since often the solution does not have the requisite regularity.”

To the best of our knowledge, no published finite difference paper takes into account the initial layer in u when proving bounds on the errors in its computed solutions. This gap in the research literature is the motivation for our work.

We shall consider the discretization of (1.1) on a mesh that is uniform in space and uniform or graded in time, where the Caputo derivative is approximated by the popular L1 scheme while a classical finite difference operator is used to approximate u_{xx} . This type of discretization has often been used in the literature, but not analyzed for problems with weak singularities at $t = 0$. Our interest here is in quantifying how the layer in the solution at $t = 0$ and the grading of the mesh influence the order of convergence of the computed solutions, measured in the discrete maximum norm. Using a precise new stability analysis of the scheme, and carefully bounding the truncation error on the graded mesh, we derive bounds on the L_∞ error in the computed solution that are shown to be sharp by numerical examples. (In [9, 10, 11], where finite element methods are used in space, the emphasis is on $L_\infty(L_2)$ rates; note that the L1 scheme is used in [9, 11].)

The structure of the paper is as follows. In section 2 the regularity of the solution u of (1.1) is investigated and the presence of an initial layer at $t = 0$ is revealed. Bounds are derived on those derivatives of the solution that are needed for the subsequent numerical analysis. In section 3 we present the discretization of (1.1) and briefly describe its convergence properties if u is smooth, i.e., if u has no layer at $t = 0$. A numerical example shows that the rate of convergence deteriorates significantly when an initial layer is present. With a view to proving bounds on the error in the computed solution in the realistic setting where u has a layer at $t = 0$, we derive a sharp new

discrete stability result for the scheme in section 4. This is used in section 5 to prove a precise convergence result for the scheme. In section 6, numerical results demonstrate the sharpness of our theoretical convergence results.

Throughout the paper we retain explicitly all dependence on the parameter T , as in some applications large values of T are used.

Notation: We use the standard Hilbert space $L_2(0, l)$ with norm $\|\cdot\|_2$ and inner product (\cdot, \cdot) . In this paper C denotes a generic constant that depends on the data of the boundary value problem (1.1) but is independent of T and of any mesh used to solve (1.1) numerically; note that C can take different values in different places.

2. Regularity of the solution to (1.1). The regularity of the solution u of (1.1) is discussed in this section. We shall see that u is smooth away from $t = 0$ but it has in general a certain singular behavior at $t = 0$.

Using separation of variables and eigenvalue expansions for the one-dimensional problem $-pg'' + cg = \lambda g$, Luchko [16] and Sakamoto and Yamamoto [19] show that one can construct an infinite series that is a classical solution $u(x, t)$ of (1.1), i.e., u satisfies the differential equation and initial-boundary conditions of (1.1) in a pointwise sense. Of course these analyses require some regularity and compatibility assumptions on the data of (1.1). We shall work in a regularity framework similar to [19], using fractional powers of the operator \mathcal{L} defined in (2.1), since it requires less smoothness of the data than the approach of [16] which uses only integer powers of \mathcal{L} .

Let $\{(\lambda_i, \psi_i) : i = 1, 2, \dots\}$ be the eigenvalues and eigenfunctions for the Sturm–Liouville two-point boundary value problem

$$(2.1) \quad \mathcal{L}\psi_i := -p\psi_i'' + c\psi_i = \lambda_i\psi_i \text{ on } (0, l), \quad \psi_i(0) = \psi_i(l) = 0,$$

where the eigenfunctions are normalized by requiring $\|\psi_i\|_2 = 1$ for all i . It is well known that $\lambda_i > 0$ for all i . A standard separation of variables technique leads formally to

$$(2.2) \quad u(x, t) = \sum_{i=1}^{\infty} \left[(\phi, \psi_i) E_{\alpha,1}(-\lambda_i t^\alpha) + J_i(t) \right] \psi_i(x)$$

(see [16, (4.29)] or [19, (2.11)]) where

$$J_i(t) := \int_{s=0}^t s^{\alpha-1} E_{\alpha,\alpha}(-\lambda_i s^\alpha) f_i(t-s) ds \quad \text{with} \quad f_i(t) := (f(\cdot, t), \psi_i(\cdot)),$$

and the generalized Mittag–Leffler function [18, section 1.2] is defined by

$$E_{\rho,\beta}(z) := \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(\rho k + \beta)}.$$

Now (2.2) defines a classical solution of (1.1), and pointwise differentiation of the terms in (2.2) yields infinite series formulas for the derivatives that we desire to bound. These results can be obtained under suitable hypotheses on the data ensuring convergence of each series; see, for example, the proof of [16, Theorem 4.1]. We shall discuss here in detail the convergence of the series for u and u_t only, as the argument for other derivatives is similar.

By [18, Theorem 1.6] there exists a constant C such that

$$(2.3) \quad |E_{\rho,\beta}(-r)| \leq \frac{C}{1+r} \text{ for } 0 < \rho < 2, \beta \in \mathbb{R}, \text{ and all } r \geq 0.$$

From the theory of sectorial operators [7, 19], the fractional power \mathcal{L}^γ of the operator \mathcal{L} is defined for each $\gamma \in \mathbb{R}$ with domain

$$D(\mathcal{L}^\gamma) := \left\{ g \in L_2(0, l) : \sum_{i=1}^{\infty} \lambda_i^{2\gamma} |(g, \psi_i)|^2 < \infty \right\}.$$

For example, $D(\mathcal{L}^{1/2}) = H_0^1(0, l)$ in Sobolev space notation. Also set

$$\|g\|_{\mathcal{L}^\gamma} := \left(\sum_{i=1}^{\infty} \lambda_i^{2\gamma} |(g, \psi_i)|^2 \right)^{1/2}.$$

Assume that $\phi \in D(\mathcal{L}^{1/2})$, and that $f(\cdot, t) \in D(\mathcal{L}^{1/2})$ for each $t \in [0, T]$ with $\|f(\cdot, t)\|_{\mathcal{L}^{1/2}} \leq C$ for some constant C independent of t . Consider the terms in (2.2): by (2.3) and $|\psi_i(x)| \leq C$ for all i and x [3, p. 335], one has

$$\begin{aligned} \sum_{i=1}^{\infty} |(\phi, \psi_i) E_{\alpha,1}(-\lambda_i t^\alpha) \psi_i(x)| &\leq C \sum_{i=1}^{\infty} |(\phi, \psi_i)| \\ &\leq C \left(\sum_{i=1}^{\infty} \frac{1}{\lambda_i} \right)^{1/2} \left(\sum_{i=1}^{\infty} \lambda_i |(\phi, \psi_i)|^2 \right)^{1/2} \\ &\leq C \end{aligned}$$

since $\phi \in D(\mathcal{L}^{1/2})$ and $\lambda_i \approx i^2$ [3, p. 415]. Furthermore, a Cauchy–Schwarz inequality and the assumption that $f(\cdot, t) \in D(\mathcal{L}^{1/2})$ for $t \in [0, T]$ give

$$(2.4) \quad \sum_{i=1}^{\infty} |f_i(t-s)| \leq \left(\sum_{i=1}^{\infty} \frac{1}{\lambda_i} \right)^{1/2} \left(\sum_{i=1}^{\infty} \lambda_i f_i^2(t-s) \right)^{1/2} \leq C \quad \text{for } 0 \leq s \leq t,$$

so

$$\sum_{i=1}^{\infty} |J_i(t) \psi_i(x)| \leq C \int_{s=0}^t s^{\alpha-1} \sum_{i=1}^{\infty} |f_i(t-s)| ds \leq C,$$

where we again used $\lambda_i \approx i^2$ and $|\psi_i(x)| \leq C$. Thus the series (2.2) is absolutely and uniformly convergent on \bar{Q} and defines a function that we call v for the moment. We have shown that

$$|v(x, t)| \leq C \quad \text{for } (x, t) \in \bar{Q}.$$

Next, differentiating (2.2) term by term with respect to t for $(x, t) \in Q$ yields

$$(2.5) \quad \begin{aligned} & - \sum_{i=1}^{\infty} \left[(\phi, \psi_i) \lambda_i t^{\alpha-1} E_{\alpha,\alpha}(-\lambda_i t^\alpha) + t^{\alpha-1} E_{\alpha,\alpha}(-\lambda_i t^\alpha) f_i(0) \right. \\ & \left. + \int_{s=0}^t s^{\alpha-1} E_{\alpha,\alpha}(-\lambda_i s^\alpha) f_i'(t-s) ds \right] \psi_i(x), \end{aligned}$$

where we used a standard formula [16, p. 154, line 2] to differentiate $E_{\alpha,1}(\cdot)$.

Assume now that (i) $\phi \in D(\mathcal{L}^{3/2})$, (ii) $f(\cdot, t) \in D(\mathcal{L}^{1/2})$ for each $t \in [0, T]$ with $\|f(\cdot, t)\|_{\mathcal{L}^{1/2}} \leq C$, and (iii) $\|f_t(\cdot, t)\|_{\mathcal{L}^{1/2}} \leq C t^{-\rho}$ for all $t \in (0, T]$ and some constant $\rho < 1$ (here C is a constant independent of t). To bound the first term in (2.5), one has

$$\sum_{i=1}^{\infty} |\lambda_i (\phi, \psi_i)| \leq \left(\sum_{i=1}^{\infty} \frac{1}{\lambda_i} \right)^{1/2} \left(\sum_{i=1}^{\infty} \lambda_i^3 |(\phi, \psi_i)|^2 \right)^{1/2} \leq C$$

by assumption (i) and $\lambda_i \approx i^2$ [3, p. 415]. Then (2.3) and $|\psi_i(x)| \leq C$ for all i and x [3, p. 335] yield

$$(2.6) \quad \sum_{i=1}^{\infty} |(\phi, \psi_i) \lambda_i t^{\alpha-1} E_{\alpha,\alpha}(-\lambda_i t^\alpha) \psi_i(x)| \leq C t^{\alpha-1}.$$

For the second term in (2.5), appealing to (2.3), (2.4) with $s = t$, and $|\psi_i(x)| \leq C$, we get

$$(2.7) \quad \sum_{i=1}^{\infty} |t^{\alpha-1} E_{\alpha,\alpha}(-\lambda_i t^\alpha) f_i(0) \psi_i(x)| \leq C t^{\alpha-1} \sum_{i=1}^{\infty} |f_i(0)| \leq C t^{\alpha-1}.$$

Finally, consider the third term in (2.5). One gets $\sum_{i=1}^{\infty} |f'_i(t-s)| \leq C(t-s)^{-\rho}$ from assumption (iii), similarly to (2.4). Hence

$$(2.8) \quad \begin{aligned} \sum_{i=1}^{\infty} \left| \psi_i(x) \int_{s=0}^t s^{\alpha-1} E_{\alpha,\alpha}(-\lambda_i s^\alpha) f'_i(t-s) ds \right| &\leq C \int_{s=0}^t s^{\alpha-1} \sum_{i=1}^{\infty} |f'_i(t-s)| ds \\ &\leq C \int_{s=0}^t s^{\alpha-1} (t-s)^{-\rho} ds \\ &\leq C t^{\alpha-1}, \end{aligned}$$

where the final integral is evaluated using Euler’s beta function [5, Theorem D.6]. Combining (2.5), (2.6), (2.7), and (2.8), we see that for each fixed $t \in (0, T]$, the series (2.5) is absolutely and uniformly convergent for $(x, t) \in [0, l] \times [\varepsilon, t]$ for each $\varepsilon > 0$, so it equals $v_t(x, t)$ on Q and we obtain

$$|v_t(x, t)| \leq C t^{\alpha-1} \text{ for } (x, t) \in Q.$$

One can carry out similar calculations to bound $v_{tt}, v_x, v_{xx}, v_{xxx}$, and v_{xxxx} on Q , where for the x derivatives one needs bounds on sums like $\sum_{i=1}^{\infty} |\psi_i(x)|^2 / \lambda_i$ from [3, p. 427]; see [16, (4.13)]. Moreover, one can show (cf. [16, 19]) that $D_t^\alpha v$ exists and v is the solution of (1.1), i.e., $v = u$. (By a maximum principle argument [15] the solution is unique.)

We summarize all this activity in the following result.

THEOREM 2.1. *Assume that $\phi \in D(\mathcal{L}^{5/2})$, $f(\cdot, t) \in D(\mathcal{L}^{5/2})$, $f_t(\cdot, t)$ and $f_{tt}(\cdot, t)$ are in $D(\mathcal{L}^{1/2})$ for each $t \in (0, T]$ with*

$$\|f(\cdot, t)\|_{\mathcal{L}^{5/2}} + \|f_t(\cdot, t)\|_{\mathcal{L}^{1/2}} + t^\rho \|f_{tt}(\cdot, t)\|_{\mathcal{L}^{1/2}} \leq C_1$$

for all $t \in (0, T]$ and some constant $\rho < 1$, where C_1 is a constant independent of t . Then (1.1) has a unique solution u that satisfies (1.1a), (1.1b), and (1.1c) pointwise, and there exists a constant C such that

$$(2.9a) \quad \left| \frac{\partial^k u}{\partial x^k}(x, t) \right| \leq C \quad \text{for } k = 0, 1, 2, 3, 4,$$

$$(2.9b) \quad \left| \frac{\partial^\ell u}{\partial t^\ell}(x, t) \right| \leq C(1 + t^{\alpha-\ell}) \quad \text{for } \ell = 0, 1, 2,$$

for all $(x, t) \in [0, l] \times (0, T]$.

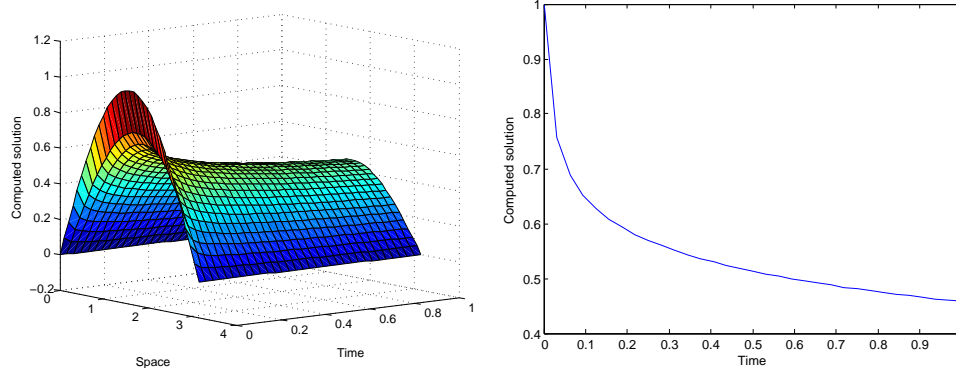


FIG. 1. Solution $v(x,t)$ of Example 2.2 when $\alpha = 0.3$, with its cross-section at $x = \pi/2$ shown in the graph on the right.

An alternative regularity analysis of (1.1) is given in [4, section 8]. It yields (2.9b), and (2.9a) for $k = 0, 1, 2$, but when $\alpha < 1/2$ it requires more regularity of the data than the analysis outlined above.

The next example shows that the bounds of (2.9) are sharp.

Example 2.2. Consider the problem $D_t^\alpha v - \partial^2 v / \partial x^2 = 0$ on $[0, \pi] \times [0, 1]$ with initial condition $v(x, 0) = \sin x$ and boundary conditions $v(0, t) = v(\pi, t) = 0$. This is a special case of (1.1) with $p = 1, c = 0$, and $l = \pi$. One can easily verify that its solution is $v(x, t) = E_{\alpha, 1}(-t^\alpha) \sin x$. Now one can compute the derivatives of v and see that they agree precisely with (2.9).

In particular, the bound (2.9b) indicates that v has a weak singularity at $t = 0$: for at least some x , one has $|v_t(x, 0)| \rightarrow \infty$ as $t \rightarrow 0^+$. Figure 1 plots the surface $v(x, t)$ and its cross section at $x = \pi/2$ for the value $\alpha = 0.3$. We observe an initial layer in v at $t = 0$, in agreement with (2.9b).

Assumption. For the remainder of this paper we shall assume that the solution u of (1.1) satisfies the bounds (2.9).

Thus in general, like Example 2.2, the solution u of (1.1) will have a weak singularity along $t = 0$. Its presence leads to significant practical and theoretical difficulties in designing and analyzing numerical methods for (1.1).

Remark 2.3. Suppose that the solution of (1.1) is smoother than indicated by (2.9b), i.e., one has instead

$$(2.10) \quad \left| \frac{\partial^\ell u}{\partial t^\ell}(x, t) \right| \leq C(1 + t^{\gamma - \ell}) \quad \text{for } \ell = 0, 1, 2 \text{ and some constant } \gamma > \alpha.$$

Then [20, Theorem 2.1] shows that the initial condition ϕ must satisfy the differential equation $-p\phi''(x) + c(x)\phi(x) = f(x, 0)$ for $x \in (0, 1)$. But the hypothesis $\phi \in D(\mathcal{L}^{1/2}) = H_0^1(0, l)$ in our discussion earlier in this section implies that $\phi(0) = \phi(1) = 0$. It follows that now *the initial condition ϕ is determined uniquely by the other data of (1.1)*, which is unacceptably restrictive.

Remark 2.4. In the numerical analysis that follows, we use no material from section 2 other than (2.9). Our results can easily be extended, for example, to the more general differential equation $D_t^\alpha u - pu_{xx} + q(x, t)u_x + c(x, t)u = f(x, t)$ if one can prove a priori that its solution satisfies (2.9). See Remark 5.4.

3. The discrete problem. The solution of the problem (1.1) is approximated by the solution of a finite difference scheme on a mesh that is uniform in space and (possibly) graded in time. Let M and N be positive integers. Set $h = l/M$ and $x_m := mh$ for $m = 0, 1, \dots, M$. Set $t_n = T(n/N)^r$ for $n = 0, 1, \dots, N$, where the constant mesh grading $r \geq 1$ is chosen by the user. If $r = 1$, then the mesh is uniform. Set $\tau_n = t_n - t_{n-1}$ for $n = 1, 2, \dots, N$. Then we have that the mesh is $\{(x_m, t_n) : m = 0, 1, \dots, M, n = 0, 1, \dots, N\}$.

The nodal approximation to the solution u computed at the mesh point (x_m, t_n) is denoted by u_m^n .

Remark 3.1. Weak singularities analogous to (2.9b) appear in the solutions to weakly singular Volterra integral equations of the second kind, and for these problems it is well known that graded meshes yield more accurate numerical solutions than uniform meshes; see [1, Chapter 6]. This is the motivation for our choice of mesh, and its efficacy will be confirmed by the theoretical and numerical results that we provide later.

The diffusion term is approximated by a standard second-order discretization:

$$u_{xx}(x_m, t_n) \approx \delta_x^2 u_m^n := \frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{h^2}.$$

The Caputo fractional derivative $D_t^\alpha u$, which can be written as

$$D_t^\alpha u(x_m, t_n) = \frac{1}{\Gamma(1-\alpha)} \sum_{k=0}^{n-1} \int_{s=t_k}^{t_{k+1}} (t_n - s)^{-\alpha} \frac{\partial u(x_m, s)}{\partial s} ds,$$

is approximated by the classical L1 approximation

$$\begin{aligned} D_N^\alpha u_m^n &:= \frac{1}{\Gamma(1-\alpha)} \sum_{k=0}^{n-1} \frac{u_m^{k+1} - u_m^k}{\tau_{k+1}} \int_{s=t_k}^{t_{k+1}} (t_n - s)^{-\alpha} ds \\ (3.1) \quad &= \frac{1}{\Gamma(2-\alpha)} \sum_{k=0}^{n-1} \frac{u_m^{k+1} - u_m^k}{\tau_{k+1}} [(t_n - t_k)^{1-\alpha} - (t_n - t_{k+1})^{1-\alpha}]. \end{aligned}$$

Thus we approximate (1.1) by the discrete problem

$$(3.2a) \quad L_{M,N} u_m^n := D_N^\alpha u_m^n - p \delta_x^2 u_m^n + c(x_m) u_m^n = f(x_m, t_n) \quad \text{for } 1 \leq m \leq M-1, 1 \leq n \leq N,$$

$$(3.2b) \quad u_0^n = 0, \quad u_M^n = 0 \quad \text{for } 0 < n \leq N,$$

$$(3.2c) \quad u_m^0 = \phi(x_m) \quad \text{for } 0 \leq m \leq M.$$

This discretization of (1.1) is standard on uniform meshes; it is considered, for example, in [14].

The pointwise bounds on the error in numerically approximating the solution of (1.1) will depend on the mesh and on the parameter α . In our subsequent analysis we shall be careful to trace the exact dependence of the order of convergence of our numerical method on α .

3.1. Analysis of discretization on uniform mesh for smooth u . Assume that $r = 1$ in section 3.1, i.e., that the mesh is uniform in space and time. Suppose

that $\partial^k u / \partial x^k \in C(\bar{Q})$ for $k = 0, \dots, 4$ and $\partial^\ell u / \partial t^\ell \in C(\bar{Q})$ for $\ell = 1, 2$. Then one can show that the discretizations of the various derivatives satisfy

$$(3.3a) \quad \frac{\partial^2 u}{\partial x^2}(x_m, t_n) = \delta_x^2 u(x_m, t_n) + O(h^2),$$

$$(3.3b) \quad D_t^\alpha u(x_m, t_n) = D_N^\alpha u(x_m, t_n) + O(\tau^{2-\alpha}).$$

The first of these is standard while (3.3b) can be found in [13].

For any mesh function $\{z_m^n\}$, set

$$\|z^n\|_\infty := \max_{0 \leq m \leq M} |z_m^n| \quad \text{and} \quad \|z\|_\infty := \max_{0 \leq n \leq N} \max_{0 \leq m \leq M} |z_m^n|.$$

The discrete stability estimate

$$(3.4) \quad \|u^n\|_\infty \leq \|u^0\|_\infty + \frac{\tau^\alpha \Gamma(2-\alpha)}{n^{1-\alpha} - (n-1)^{1-\alpha}} \|f\|_\infty \quad \text{for } n = 1, 2, \dots, N$$

is derived in [14, Theorem 2] for the case $c \equiv 0$. This result is easily extended to $c \geq 0$.

It is now straightforward to prove the following result, which is a slight modification of [14, Theorem 4].

THEOREM 3.2. *Under the smoothness assumptions made above for u on \bar{Q} , the solution of (3.2) satisfies*

$$(3.5) \quad \max_{(x_m, t_n) \in \bar{Q}} |u(x_m, t_n) - u_m^n| \leq C(h^2 + \tau^{2-\alpha})$$

for some constant C .

We give a numerical example to demonstrate that the $O(\tau^{2-\alpha})$ component of the error in (3.5) is sharp (the $O(h^2)$ spatial error in this problem is routine; it is the temporal error due to the approximation of the fractional derivative that is of interest).

In the numerical examples here and in section 6, each box of the table of results contains the maximum nodal error

$$E^{M,N} := \max_{(x_m, t_n) \in \bar{Q}} |u(x_m, t_n) - u_m^n|$$

and the corresponding rate of convergence, which is computed using the standard formula

$$\text{rate}_{M,N} := \log_2 \left(\frac{E^{M,N}}{E^{2M,2N}} \right).$$

Example 3.3. In (1.1) take $p = 1, c = 0, l = \pi$, and $T = 1$. Choose $u(x, t) = t^3 \sin x$, so $\phi(x) \equiv 0$. Then $f(x, t) = t^3 \sin x + [\Gamma(4)/\Gamma(4-\alpha)]t^{3-\alpha} \sin x$. Note that while $u \in C^\infty(\bar{Q})$, one has $f \notin C^3(\bar{Q})$.

Numerical results are given in Table 1. They indicate that the order of convergence of the method is $\tau^{2-\alpha}$, which of course dominates the $O(h^2)$ component of the error in Theorem 3.2 since $M = N$ throughout this table. We observe that, as the parameter α increases, the magnitudes of the maximum errors increase.

TABLE 1
Maximum nodal errors and orders of convergence for Example 3.3.

	N=M=64	N=M=128	N=M=256	N=M=512	N=M=1024	N=M=2048	N=M=4096
$\alpha = 0.2$	2.350E-4 1.819	6.660E-5 1.817	1.891E-5 1.815	5.374E-6 1.813	1.529E-6 1.812	4.357E-7 1.810	1.242E-7
$\alpha = 0.4$	6.969E-4 1.603	2.294E-4 1.603	7.552E-5 1.602	2.487E-5 1.602	8.195E-6 1.601	2.701E-6 1.601	8.902E-7
$\alpha = 0.6$	2.122E-3 1.397	8.058E-4 1.398	3.057E-4 1.399	1.159E-4 1.399	4.395E-5 1.400	1.666E-5 1.400	6.313E-6
$\alpha = 0.8$	6.217E-3 1.197	2.712E-3 1.198	1.182E-3 1.199	5.148E-4 1.199	2.242E-4 1.200	9.761E-5 1.200	4.249E-5

TABLE 2
Maximum nodal errors and orders of convergence for Example 2.2.

	N=M=64	N=M=128	N=M=256	N=M=512	N=M=1024	N=M=2048	N=M=4096
$\alpha = 0.2$	3.704E-2 0.099	3.458E-2 0.109	3.207E-2 0.118	2.955E-2 0.127	2.706E-2 0.135	2.465E-2 0.142	2.233E-2
$\alpha = 0.4$	3.562E-2 0.326	2.841E-2 0.343	2.240E-2 0.356	1.749E-2 0.367	1.357E-2 0.375	1.046E-2 0.381	8.038E-3
$\alpha = 0.6$	1.827E-2 0.577	1.225E-2 0.585	8.166E-3 0.590	5.424E-3 0.594	3.594E-3 0.596	2.378E-3 0.597	1.572E-3
$\alpha = 0.8$	6.342E-3 0.811	3.616E-3 0.807	2.068E-3 0.803	1.185E-3 0.802	6.794E-4 0.801	3.898E-4 0.801	2.238E-4

3.2. Accuracy of discretization on uniform mesh for typical u . Assume that $r = 1$ in section 3.2. The results of section 3.1 show that Theorem 3.2 is sharp when the solution u of (1.1) is sufficiently smooth. But typical solutions of (1.1) are nonsmooth, as we saw in section 2. Thus we now test the finite difference scheme (3.2) on Example 2.2 to see how well it performs on a problem that is truly representative of solutions of (1.1).

Table 2 displays the maximum nodal errors and orders of convergence for Example 2.2. These indicate that, assuming one has $O(h^2)$ convergence in the spatial variable, the method converges with order τ^α . [For low orders of convergence (e.g., $\alpha = 0.2$), a very fine mesh would be required before the asymptotic rate of convergence would be attained; consequently, when $\alpha = 0.2$ in Table 2, one sees slow convergence of the rates to the expected asymptotic rate of 0.2.] This rate of convergence τ^α is clearly lower than the $O(\tau^{2-\alpha})$ that was proved in Theorem 3.2. Now, unlike Table 1, the rates of convergence increase and the error decreases as α increases, i.e., as (1.1a) comes closer to a classical parabolic partial differential equation.

We shall devise a convergence analysis for our graded-mesh scheme applied to solutions u of (1.1) that satisfy the derivative bounds (2.9). This analysis will prove that on a uniform mesh one attains the $O(h^2 + \tau^\alpha)$ rate of convergence that we observed in Table 2, and will also predict the optimal grading r that should be used in our mesh to attain the highest order of convergence; see Theorem 5.3.

4. Stability of the scheme. We now return to arbitrarily graded meshes in time, i.e., $r \geq 1$. The first task is to make the discrete stability result (3.4) more precise by showing that $\|u^n\|_\infty$ depends on $\|u^0\|_\infty$ and on $\|f^k\|_\infty$ for $k = 1, 2, \dots, n$, with the dependence on $\|f^k\|_\infty$ strongest when k is near n .

Rewrite the discretization (3.1) of the Caputo derivative as

$$D_N^\alpha u_m^n = \frac{d_{n,1}}{\Gamma(2-\alpha)} u_m^n - \frac{d_{n,n}}{\Gamma(2-\alpha)} u_m^0 + \frac{1}{\Gamma(2-\alpha)} \sum_{k=1}^{n-1} u_m^{n-k} [d_{n,k+1} - d_{n,k}],$$

where

$$(4.1) \quad d_{n,k} := \frac{(t_n - t_{n-k})^{1-\alpha} - (t_n - t_{n-k+1})^{1-\alpha}}{\tau_{n-k+1}}.$$

In particular $d_{n,1} = \tau_n^{-\alpha}$.

Using the mean value theorem one can prove easily that

$$(4.2) \quad d_{n,k+1} \leq d_{n,k},$$

and one also has

$$(4.3) \quad (1-\alpha)(t_n - t_{n-k})^{-\alpha} \leq d_{n,k} \leq (1-\alpha)(t_n - t_{n-k+1})^{-\alpha}.$$

Our scheme is now

$$\left[\frac{d_{n,1}}{\Gamma(2-\alpha)} + \frac{2p}{h^2} + c(x_m) \right] u_m^n = \frac{p}{h^2} u_{m+1}^n + \frac{p}{h^2} u_{m-1}^n + f(x_m, t_n) + \frac{1}{\Gamma(2-\alpha)} \left[d_{n,n} u_m^0 + \sum_{k=1}^{n-1} (d_{n,k} - d_{n,k+1}) u_m^{n-k} \right]$$

for $m = 1, 2, \dots, M-1$ and $n = 1, 2, \dots, N$.

Our stability result will be presented in a general framework. Assume that

$$(4.4) \quad L_{M,N} v_m^n = g_m^n \text{ for } 1 \leq m \leq M-1, 1 \leq n \leq N,$$

with $v_0^n = v_N^n = 0$ for $0 < n \leq N$ and v_m^0 given for $0 \leq m \leq M$.

The next result is the fundamental stability bound for (3.2). It is related to [14, Lemma 3].

LEMMA 4.1. *The solution of the discrete problem (4.4) satisfies*

$$(4.5) \quad \|v^n\|_\infty \leq \tau_n^\alpha \left[\Gamma(2-\alpha) \|g^n\|_\infty + d_{n,n} \|v^0\|_\infty + \sum_{k=1}^{n-1} (d_{n,k} - d_{n,k+1}) \|v^{n-k}\|_\infty \right]$$

for $n = 1, 2, \dots, N$.

Proof. Fix $n \in \{1, 2, \dots, N\}$. Choose j_0 such that $|v_{j_0}^n| = \|v^n\|_\infty$. The equation associated with the mesh point (x_{j_0}, t_n) is

$$\left[\frac{d_{n,1}}{\Gamma(2-\alpha)} + \frac{2p}{h^2} + c(x_{j_0}) \right] v_{j_0}^n = \frac{p}{h^2} v_{j_0+1}^n + \frac{p}{h^2} v_{j_0-1}^n + g_{j_0}^n + \frac{1}{\Gamma(2-\alpha)} \left[d_{n,n} v_{j_0}^0 + \sum_{k=1}^{n-1} (d_{n,k} - d_{n,k+1}) v_{j_0}^{n-k} \right].$$

Hence, by $c \geq 0$ and the choice of j_0 , one obtains

$$\begin{aligned} & \left[\frac{d_{n,1}}{\Gamma(2-\alpha)} + \frac{2p}{h^2} \right] \|v^n\|_\infty \\ & \leq \frac{2p}{h^2} \|v^n\|_\infty + \left| g_{j_0}^n + \frac{1}{\Gamma(2-\alpha)} \left[d_{n,n}v_{j_0}^0 + \sum_{k=1}^{n-1} (d_{n,k} - d_{n,k+1})v_{j_0}^{n-k} \right] \right|, \end{aligned}$$

which is equivalent to

$$\frac{d_{n,1}}{\Gamma(2-\alpha)} \|v^n\|_\infty \leq \left| g_{j_0}^n + \frac{1}{\Gamma(2-\alpha)} \left[d_{n,n}v_{j_0}^0 + \sum_{k=1}^{n-1} (d_{n,k} - d_{n,k+1})v_{j_0}^{n-k} \right] \right|.$$

The desired result now follows easily on recalling the inequality (4.2). □

Lemma 4.1 will now be used in an inductive argument to give a weighted bound for $\|v^n\|_\infty$ in terms of the given data $\|v^0\|_\infty$ and $\|g^j\|_\infty$ for $j = 1, 2, \dots, n$.

Define the real numbers $\theta_{n,j}$, for $n = 1, 2, \dots, N$ and $j = 1, 2, \dots, n - 1$ by

$$(4.6) \quad \theta_{n,n} = 1, \quad \theta_{n,j} = \sum_{k=1}^{n-j} \tau_{n-k}^\alpha (d_{n,k} - d_{n,k+1}) \theta_{n-k,j}.$$

Observe that (4.2) implies $\theta_{n,j} > 0$ for all n, j .

LEMMA 4.2. *The solution of the discrete problem (4.4) satisfies*

$$(4.7) \quad \|v^n\|_\infty \leq \|v^0\|_\infty + \tau_n^\alpha \Gamma(2-\alpha) \sum_{j=1}^n \theta_{n,j} \|g^j\|_\infty$$

for $n = 1, 2, \dots, N$.

Proof. Use induction on n . The case $n = 1$ of (4.7) is

$$\|v^1\|_\infty \leq \tau_1^\alpha d_{1,1} \|v^0\|_\infty + \tau_1^\alpha \Gamma(2-\alpha) \theta_{1,1} \|g^1\|_\infty$$

which is identical to (4.5).

Fix $n \in \{2, 3, \dots, N\}$. Assume that (4.7) is valid for $k = 1, 2, \dots, n-1$. Then (4.5) and the inductive hypothesis yield

$$\begin{aligned} \|v^n\|_\infty & \leq \tau_n^\alpha \left\{ \Gamma(2-\alpha) \|g^n\|_\infty + d_{n,n} \|v^0\|_\infty \right. \\ & \quad \left. + \sum_{k=1}^{n-1} (d_{n,k} - d_{n,k+1}) \left[\|v^0\|_\infty + \tau_{n-k}^\alpha \Gamma(2-\alpha) \sum_{j=1}^{n-k} \theta_{n-k,j} \|g^j\|_\infty \right] \right\} \\ & = \tau_n^\alpha \left\{ \Gamma(2-\alpha) \|g^n\|_\infty + d_{n,1} \|v^0\|_\infty \right. \\ & \quad \left. + \Gamma(2-\alpha) \sum_{k=1}^{n-1} \left[\tau_{n-k}^\alpha (d_{n,k} - d_{n,k+1}) \sum_{j=1}^{n-k} \theta_{n-k,j} \|g^j\|_\infty \right] \right\} \end{aligned}$$

$$\begin{aligned}
 &= \tau_n^\alpha \left\{ \Gamma(2 - \alpha) \|g^n\|_\infty + d_{n,1} \|v^0\|_\infty \right. \\
 &\quad \left. + \Gamma(2 - \alpha) \sum_{j=1}^{n-1} \|g^j\|_\infty \left[\sum_{k=1}^{n-j} \tau_{n-k}^\alpha (d_{n,k} - d_{n,k+1}) \theta_{n-k,j} \right] \right\} \\
 &= \|v^0\|_\infty + \tau_n^\alpha \Gamma(2 - \alpha) \left\{ \sum_{j=1}^n \|g^j\|_\infty \left[\sum_{k=1}^{n-j} \tau_{n-k}^\alpha (d_{n,k} - d_{n,k+1}) \theta_{n-k,j} \right] \right\} \\
 &= \|v^0\|_\infty + \tau_n^\alpha \Gamma(2 - \alpha) \sum_{j=1}^n \theta_{n,j} \|g^j\|_\infty,
 \end{aligned}$$

where we used telescoping to handle $\|v^0\|_\infty$ and (4.6) to handle the $\theta_{n-k,j}$. Thus we have proved (4.7) for $k = n$. By the principle of induction, the lemma is proved. \square

The next result bounds a weighted sum of the $\theta_{n,j}$ that will be needed in our error analysis.

LEMMA 4.3. *Let the parameter β satisfy $\beta \leq r\alpha$. Then for $n = 1, 2, \dots, N$, one has*

$$(4.8) \quad \tau_n^\alpha \sum_{j=1}^n j^{-\beta} \theta_{n,j} \leq \frac{T^\alpha N^{-\beta}}{1 - \alpha}.$$

Proof. Use induction on n . When $n = 1$, then

$$\tau_1^\alpha \sum_{j=1}^1 j^{-\beta} \theta_{1,j} = T^\alpha N^{-r\alpha} \leq T^\alpha N^{-\beta}$$

by (4.6), $\tau_1 = TN^{-r}$, and $\beta \leq r\alpha$. Thus (4.8) is true for $n = 1$.

Next, assume that (4.8) is true for $n = 1, 2, \dots, \ell - 1$, where $\ell \in \{2, 3, \dots, N\}$. We want to prove (4.8) for $n = \ell$. Invoking (4.6) and interchanging the order of summation,

$$\begin{aligned}
 \tau_\ell^\alpha \sum_{j=1}^\ell j^{-\beta} \theta_{\ell,j} &= \tau_\ell^\alpha \ell^{-\beta} \theta_{\ell,\ell} + \tau_\ell^\alpha \sum_{j=1}^{\ell-1} j^{-\beta} \sum_{k=1}^{\ell-j} \tau_{\ell-k}^\alpha (d_{\ell,k} - d_{\ell,k+1}) \theta_{\ell-k,j} \\
 &= \tau_\ell^\alpha \ell^{-\beta} + \tau_\ell^\alpha \sum_{k=1}^{\ell-1} (d_{\ell,k} - d_{\ell,k+1}) \tau_{\ell-k}^\alpha \sum_{j=1}^{\ell-k} j^{-\beta} \theta_{\ell-k,j} \\
 &\leq \tau_\ell^\alpha \ell^{-\beta} + \tau_\ell^\alpha \sum_{k=1}^{\ell-1} (d_{\ell,k} - d_{\ell,k+1}) \frac{T^\alpha N^{-\beta}}{1 - \alpha}
 \end{aligned}$$

by the inductive hypothesis. Now telescoping and the definition of $d_{\ell,1}$ yield

$$\begin{aligned}
 \tau_\ell^\alpha \sum_{j=1}^\ell j^{-\beta} \theta_{\ell,j} &\leq \tau_\ell^\alpha \ell^{-\beta} + \tau_\ell^\alpha (d_{\ell,1} - d_{\ell,\ell}) \frac{T^\alpha N^{-\beta}}{1 - \alpha} \\
 &= \frac{T^\alpha N^{-\beta}}{1 - \alpha} + \tau_\ell^\alpha \left[\ell^{-\beta} - d_{\ell,\ell} \frac{T^\alpha N^{-\beta}}{1 - \alpha} \right].
 \end{aligned}$$

To complete the inductive step, we show that $[\dots] \leq 0$. For

$$\begin{aligned} d_{\ell,\ell} \frac{T^\alpha N^{-\beta}}{1-\alpha} &= \frac{t_\ell^{1-\alpha} - (t_\ell - t_1)^{1-\alpha}}{\tau_1} \cdot \frac{T^\alpha N^{-\beta}}{1-\alpha} \\ &= \frac{T^{1-\alpha} [\ell^{r(1-\alpha)} - (\ell^r - 1)^{1-\alpha}]}{N^{r(1-\alpha)} T N^{-r}} \cdot \frac{T^\alpha N^{-\beta}}{1-\alpha} \\ &= \frac{[\ell^{r(1-\alpha)} - (\ell^r - 1)^{1-\alpha}] N^{r\alpha-\beta}}{1-\alpha} \\ &\geq (\ell^r)^{-\alpha} N^{r\alpha-\beta} \\ &= \ell^{-\beta} (N/\ell)^{r\alpha-\beta} \\ &\geq \ell^{-\beta}, \end{aligned}$$

where we used the mean value theorem, $N \geq \ell$, and $r\alpha \geq \beta$. Thus (4.8) holds true for $n = \ell$.

By the principle of induction, (4.8) is true for all n . □

If the mesh is uniform (i.e., $r = 1$ and $\tau_n = \tau$ for all n), Lemmas 4.2 and 4.3 (with $\beta = 0$) imply the less precise result

$$\|v^n\|_\infty \leq \|v^0\|_\infty + \frac{n^\alpha \tau^\alpha \Gamma(2-\alpha)}{1-\alpha} \|g\|_\infty,$$

which is essentially equivalent to the old stability result (3.4).

5. Error analysis of finite difference scheme. This analysis of (3.2) on a general graded mesh proceeds by bounding the truncation error then applying the discrete stability result of Lemma 4.2.

The following bound is needed in several places:

$$(5.1) \quad \tau_{k+1} = T \left(\frac{k+1}{N} \right)^r - T \left(\frac{k}{N} \right)^r \leq CTN^{-r} k^{r-1}$$

for $k = 0, 1, \dots, N - 1$.

5.1. Truncation error for diffusion term. A standard truncation error estimate using (2.9a) yields

$$(5.2) \quad \frac{\partial^2 u}{\partial x^2}(x_m, t_n) = \delta_x^2 u(x_m, t_n) + O(h^2)$$

for all mesh points $(x_m, t_n) \in Q$.

Remark 5.1. If one has the bound (2.9a) only for $k = 0, 1, 2, 3$, then (5.2) should be replaced by $\partial^2 u(x_m, t_n)/\partial x^2 = \delta_x^2 u(x_m, t_n) + O(h)$. In the subsequent analysis, replace $O(h^2)$ everywhere by $O(h)$; no other change is needed.

5.2. Truncation error for Caputo derivative term. $m \in \{1, 2, \dots, M - 1\}$ and $n \in \{1, 2, \dots, N\}$ are fixed. Then

$$D_N^\alpha u(x_m, t_n) - D_t^\alpha u(x_m, t_n) = \sum_{k=0}^{n-1} T_{nk}$$

with

$$(5.3) \quad T_{nk} := \frac{1}{\Gamma(1-\alpha)} \int_{s=t_k}^{t_{k+1}} (t_n - s)^{-\alpha} \left[\frac{u(x_m, t_{k+1}) - u(x_m, t_k)}{\tau_{k+1}} - \frac{\partial u}{\partial s}(x_m, s) \right] ds.$$

Although T_{nk} depends on m , the notation ignores this relationship since all our bounds of T_{nk} are based on (2.9) which is independent of x .

A triangle inequality yields the bound

$$(5.4) \quad |D_N^\alpha u(x_m, t_n) - D_t^\alpha u(x_m, t_n)| \leq \sum_{k=0}^{n-1} |T_{nk}|.$$

For $k+1 < n$, an integration by parts yields

$$(5.5) \quad T_{nk} = \frac{-\alpha}{\Gamma(1-\alpha)} \int_{s=t_k}^{t_{k+1}} (t_n - s)^{-\alpha-1} (\phi - \psi)(x_m, s) ds,$$

where

$$\begin{aligned} \phi(x_m, s) &:= \left[\frac{u(x_m, t_{k+1}) - u(x_m, t_k)}{\tau_{k+1}} \right] (s - t_k), \\ \psi(x_m, s) &:= u(x_m, s) - u(x_m, t_k). \end{aligned}$$

Then the mean value theorem for integrals gives

$$T_{nk} = \frac{-\alpha}{\Gamma(1-\alpha)} (\phi - \psi)(x_m, \gamma_1) \int_{s=t_k}^{t_{k+1}} (t_n - s)^{-\alpha-1} ds$$

for some $\gamma_1 \in (t_k, t_{k+1})$. But

$$\begin{aligned} (\phi - \psi)(x_m, \gamma_1) &= \left[\frac{u(x_m, t_{k+1}) - u(x_m, t_k)}{\tau_{k+1}} \right] (\gamma_1 - t_k) - [u(x_m, \gamma_1) - u(x_m, t_k)] \\ &= (\gamma_1 - t_k) [u_t(x_m, \gamma_2) - u_t(x_m, \gamma_3)] \\ &= (\gamma_1 - t_k) (\gamma_2 - \gamma_3) u_{tt}(x_m, \gamma_4) \end{aligned}$$

for some $\gamma_2, \gamma_3, \gamma_4 \in (t_k, t_{k+1})$. Thus for $1 \leq k < n-1$,

$$(5.6) \quad |T_{nk}| \leq C \tau_{k+1}^2 \left(\max_{s \in [t_k, t_{k+1}]} |u_{tt}(x_m, s)| \right) \int_{s=t_k}^{t_{k+1}} (t_n - s)^{-\alpha-1} ds.$$

LEMMA 5.2. *There exists a constant C such that for all $(x_m, t_n) \in Q$ one has*

$$|D_N^\alpha u(x_m, t_n) - D_t^\alpha u(x_m, t_n)| \leq C n^{-\min\{2-\alpha, r\alpha\}}.$$

Proof. We use (5.4). For $1 \leq k < n-1$, by (5.6) and (2.9b) one has

$$\begin{aligned} |T_{nk}| &\leq C \tau_{k+1}^2 t_k^{\alpha-2} \tau_{k+1} (t_n - t_{k+1})^{-\alpha-1} \\ &= C \tau_{k+1}^3 T^{\alpha-2} \left(\frac{k}{N} \right)^{r(\alpha-2)} T^{-\alpha-1} \left[\left(\frac{n}{N} \right)^r - \left(\frac{k+1}{N} \right)^r \right]^{-\alpha-1} \\ &= C \tau_{k+1}^3 T^{-3} N^{3r} k^{r(\alpha-2)} [n^r - (k+1)^r]^{-\alpha-1}. \end{aligned}$$

Now (5.1) yields

$$(5.7) \quad |T_{nk}| \leq C k^{3(r-1)} k^{r(\alpha-2)} [n^r - (k+1)^r]^{-\alpha-1} = C k^{r(\alpha+1)-3} [n^r - (k+1)^r]^{-\alpha-1}$$

for $1 \leq k < n-1$. Consequently,

$$(5.8) \quad \sum_{k=1}^{\lceil n/2 \rceil - 1} |T_{nk}| \leq C \sum_{k=1}^{\lceil n/2 \rceil - 1} k^{r(\alpha+1)-3} n^{r(-\alpha-1)} = C n^{-r(\alpha+1)} \sum_{k=1}^{\lceil n/2 \rceil - 1} k^{r(\alpha+1)-3}.$$

Hence, well-known convergence results for series of this type give

$$(5.9) \quad \sum_{k=1}^{\lceil n/2 \rceil - 1} |T_{nk}| \leq \begin{cases} Cn^{-r(\alpha+1)} & \text{if } r(\alpha+1) < 2, \\ Cn^{-2} \ln n & \text{if } r(\alpha+1) = 2, \\ Cn^{-2} & \text{if } r(\alpha+1) > 2. \end{cases}$$

For $\lceil n/2 \rceil \leq k < n - 1$, start from (5.6) and invoke (5.1):

$$\begin{aligned} |T_{nk}| &\leq C\tau_{k+1}^2 \left(\max_{s \in [t_k, t_{k+1}]} |u_{tt}(x_m, s)| \right) \int_{s=t_k}^{t_{k+1}} (t_n - s)^{-\alpha-1} ds \\ &\leq C(TN^{-r}k^{r-1})^2 t_k^{\alpha-2} \int_{s=t_k}^{t_{k+1}} (t_n - s)^{-\alpha-1} ds \\ &\leq CT^2 N^{-2r} n^{2(r-1)} t_n^{\alpha-2} \int_{s=t_k}^{t_{k+1}} (t_n - s)^{-\alpha-1} ds \\ &= CT^\alpha N^{-2r} n^{2(r-1)} \left(\frac{n}{N}\right)^{r(\alpha-2)} \int_{s=t_k}^{t_{k+1}} (t_n - s)^{-\alpha-1} ds \\ &= CT^\alpha N^{-r\alpha} n^{r\alpha-2} \int_{s=t_k}^{t_{k+1}} (t_n - s)^{-\alpha-1} ds. \end{aligned}$$

Hence

$$(5.10) \quad \begin{aligned} \sum_{k=\lceil n/2 \rceil}^{n-2} |T_{nk}| &\leq CT^\alpha N^{-r\alpha} n^{r\alpha-2} \int_{s=t_{\lceil n/2 \rceil}}^{t_{n-1}} (t_n - s)^{-\alpha-1} ds \\ &\leq CT^\alpha N^{-r\alpha} n^{r\alpha-2} (t_n - t_{n-1})^{-\alpha} \\ &\leq CT^\alpha N^{-r\alpha} n^{r\alpha-2} (TN^{-r}n^{r-1})^{-\alpha} \\ &= Cn^{-(2-\alpha)}, \end{aligned}$$

where we evaluated the integral, discarded a negative term, then invoked (5.1).

In the sum $\sum_{k=0}^{n-1} |T_{nk}|$ of (5.4), it remains to bound T_{n0} and $T_{n,n-1}$. If $n = 1$, then

$$T_{10} = \frac{\tau_1^{-\alpha}}{\Gamma(2-\alpha)} [u(x_m, t_1) - u(x_m, t_0)] - \frac{1}{\Gamma(1-\alpha)} \int_{s=0}^{t_1} (t_1 - s)^{-\alpha} \frac{\partial u}{\partial s}(x_m, s) ds.$$

We derive estimates for each of these terms. The first term satisfies

$$\begin{aligned} \left| \frac{\tau_1^{-\alpha}}{\Gamma(2-\alpha)} [u(x_m, t_1) - u(x_m, t_0)] \right| &= \left| \frac{\tau_1^{-\alpha}}{\Gamma(2-\alpha)} \int_{s=0}^{\tau_1} \frac{\partial u}{\partial s}(x_m, s) ds \right| \\ &\leq \frac{\tau_1^{-\alpha}}{\Gamma(2-\alpha)} \int_{s=0}^{\tau_1} \left| \frac{\partial u}{\partial s}(x_m, s) \right| ds \\ &\leq C \frac{\tau_1^{-\alpha}}{\Gamma(2-\alpha)} \int_{s=0}^{\tau_1} s^{\alpha-1} ds \\ &= C \frac{\tau_1^{-\alpha}}{\Gamma(2-\alpha)} \frac{\tau_1^\alpha}{\alpha} \\ &\leq C. \end{aligned}$$

The second term in T_{10} satisfies

$$\begin{aligned} \left| \frac{1}{\Gamma(1-\alpha)} \int_{s=0}^{t_1} (t_1-s)^{-\alpha} \frac{\partial u}{\partial s}(x_m, s) ds \right| &\leq \frac{1}{\Gamma(1-\alpha)} \int_{s=0}^{t_1} (t_1-s)^{-\alpha} \left| \frac{\partial u}{\partial s}(x_m, s) \right| ds \\ &\leq \frac{C}{\Gamma(1-\alpha)} \int_{s=0}^{t_1} (t_1-s)^{-\alpha} s^{\alpha-1} ds \\ &= \frac{C}{\Gamma(1-\alpha)} B(\alpha, 1-\alpha) \\ &\leq C. \end{aligned}$$

Thus,

$$(5.11) \quad |T_{10}| \leq C.$$

Assume now that $n > 1$. The first term in T_{n0} can be bounded by

$$\begin{aligned} &\left| \frac{\tau_1^{-1}}{\Gamma(2-\alpha)} [u(x_m, t_1) - u(x_m, t_0)] [t_n^{1-\alpha} - (t_n - t_1)^{1-\alpha}] \right| \\ &\leq \frac{\tau_1^{-1}}{\Gamma(2-\alpha)} [t_n^{1-\alpha} - (t_n - t_1)^{1-\alpha}] \int_{s=0}^{t_1} \left| \frac{\partial u}{\partial s}(x_m, s) \right| ds \\ &\leq \frac{C\tau_1^{-1}}{\Gamma(2-\alpha)} [t_n^{1-\alpha} - (t_n - t_1)^{1-\alpha}] \int_{s=0}^{t_1} s^{\alpha-1} ds \\ &= \frac{C\tau_1^{-1}}{\Gamma(2-\alpha)} [t_n^{1-\alpha} - (t_n - t_1)^{1-\alpha}] \frac{t_1^\alpha}{\alpha} \\ &\leq C [t_n^{1-\alpha} - (t_n - t_1)^{1-\alpha}] t_1^{\alpha-1} \\ &\leq C \left(\frac{t_n - t_1}{t_1} \right)^{-\alpha}, \end{aligned}$$

where we used the mean value theorem in the final inequality. The second term is bounded by

$$\begin{aligned} \left| \frac{1}{\Gamma(1-\alpha)} \int_{s=0}^{t_1} (t_n-s)^{-\alpha} \frac{\partial u}{\partial s}(x_m, s) ds \right| &\leq \frac{(t_n - t_1)^{-\alpha}}{\Gamma(1-\alpha)} \int_{s=0}^{t_1} \left| \frac{\partial u}{\partial s}(x_m, s) \right| ds \\ &\leq \frac{C(t_n - t_1)^{-\alpha}}{\Gamma(1-\alpha)} \int_{s=0}^{t_1} s^{\alpha-1} ds \\ &= \frac{C(t_n - t_1)^{-\alpha}}{\Gamma(1-\alpha)} t_1^\alpha \\ &= C \left(\frac{t_n - t_1}{t_1} \right)^{-\alpha}. \end{aligned}$$

Hence

$$(5.12) \quad |T_{n0}| \leq C \left(\frac{t_n - t_1}{t_1} \right)^{-\alpha} \leq Cn^{-r\alpha} \quad \text{for } n > 1.$$

Combining (5.11) and (5.12), we have

$$(5.13) \quad |T_{n0}| \leq Cn^{-r\alpha} \quad \text{for } n \geq 1.$$

Finally, we consider $T_{n,n-1}$ for $n > 1$ (the case $n = 1$ is T_{10} , which has been dealt with already). From (5.3), for some $\gamma_5, \gamma_6, \gamma_7 \in (t_{n-1}, t_n)$ one gets

$$\begin{aligned}
 |T_{n,n-1}| &= \frac{1}{\Gamma(1-\alpha)} \left| \frac{\partial u}{\partial t}(x_m, \gamma_5) - \frac{\partial u}{\partial t}(x_m, \gamma_6) \right| \int_{s=t_{n-1}}^{t_n} (t_n - s)^{-\alpha} ds \\
 &\leq \frac{\tau_n}{\Gamma(2-\alpha)} \left| \frac{\partial^2 u}{\partial t^2}(x_m, \gamma_7) \right| (t_n - t_{n-1})^{1-\alpha} \\
 &\leq C \tau_n^{2-\alpha} t_{n-1}^{\alpha-2} \\
 &\leq C [TN^{-r}(n-1)^{r-1}]^{2-\alpha} T^{\alpha-2} \left(\frac{n-1}{N}\right)^{r(\alpha-2)} \\
 &= C(n-1)^{-(2-\alpha)} \\
 (5.14) \quad &\leq Cn^{-(2-\alpha)},
 \end{aligned}$$

where we used (2.9b) and (5.1).

To finish the proof, we have to combine the bounds (5.9), (5.10), (5.13), and (5.14) in (5.4). Now $n^{-r\alpha} \geq Cn^{-r(\alpha+1)} \ln n$ so (5.13) dominates (5.9) if $r(\alpha + 1) \leq 2$; on the other hand, if $r(\alpha + 1) > 2$, then $n^{-(2-\alpha)} \geq n^{-2}$ implies that (5.14) dominates (5.9). Thus (5.9) can be ignored, and it is an easy matter to combine the other bounds to obtain the desired result. \square

5.3. Convergence of the scheme. We can now prove our main convergence result.

THEOREM 5.3. *The solution u_m^n of the scheme (3.2) satisfies*

$$(5.15) \quad \max_{(x_m, t_n) \in \bar{Q}} |u(x_m, t_n) - u_m^n| \leq C \left(h^2 + T^\alpha N^{-\min\{2-\alpha, r\alpha\}} \right)$$

for some constant C .

Proof. Fix $(x_m, t_n) \in \bar{Q}$. The truncation error of (3.2) at (x_m, t_n) is

$$|\chi_m^n| := |D_M^\alpha [u(x_m, t_n) - u_m^n] - \delta_x^2 [u(x_m, t_n) - u_m^n]| \leq C \left(h^2 + n^{-\min\{2-\alpha, r\alpha\}} \right)$$

by (5.2) and Lemma 5.2. Now $L_{M,N}[u(x_m, t_n) - u_m^n] = \chi_m^n$, so invoking Lemma 4.2 with $v_m^n := u(x_m, t_n) - u_m^n$ (observe that the initial condition $v_m^0 = 0$ for all m) we obtain

$$\begin{aligned}
 \max_{(x_m, t_n) \in \bar{Q}} |u(x_m, t_n) - u_m^n| &\leq C \tau_n^\alpha \Gamma(2-\alpha) \sum_{j=1}^n \theta_{n,j} \|\chi^j\|_\infty \\
 &\leq C \tau_n^\alpha \Gamma(2-\alpha) \sum_{j=1}^n \theta_{n,j} \left(h^2 + j^{-\min\{2-\alpha, r\alpha\}} \right) \\
 &\leq CT^\alpha \left(h^2 + N^{-\min\{2-\alpha, r\alpha\}} \right)
 \end{aligned}$$

by Lemma 4.3 with $\beta = 0$ for the h^2 term and $\beta = \min\{2 - \alpha, r\alpha\}$ for the term involving $N^{-\min\{2-\alpha, r\alpha\}}$. \square

Remark 5.4 (more general differential equation). One can prove a convergence result similar to Theorem 5.3 if (1.1a) is replaced by the differential equation

$$D_t^\alpha u - p \frac{\partial^2 u}{\partial x^2} + q(x, t) \frac{\partial u}{\partial x} + c(x, t)u = f(x, t),$$

where a convection term q appears and q, c can be functions of t as well as x , provided that the solution u still satisfies the a priori bounds (2.9). The only essential modification needed in the analysis is the nonrestrictive assumption that $l\|q\|_\infty \leq 2pM$, to ensure that Lemma 4.1 remains valid.

Remark 5.5 (weaker hypothesis on temporal derivatives of u). Suppose that the bound

$$\left| \frac{\partial^\ell u}{\partial t^\ell}(x, t) \right| \leq C(1 + t^{\alpha-\ell}) \quad \text{for } \ell = 0, 1, 2$$

of (2.9b) is replaced by the more general hypothesis that

$$(5.16) \quad \left| \frac{\partial^\ell u}{\partial t^\ell}(x, t) \right| \leq C(1 + t^{\gamma-\ell}) \quad \text{for } \ell = 0, 1, 2,$$

where $\gamma \in (0, \alpha]$ is fixed. This modification has no effect on section 4, and in section 5 an inspection shows that only algebraic changes are needed: inequality (5.9) becomes

$$(5.17) \quad \sum_{k=1}^{\lceil n/2 \rceil - 1} |T_{nk}| \leq \begin{cases} Cn^{-r(\alpha+1)}(T^{-1}N^r)^{\alpha-\gamma} & \text{if } r(\gamma+1) < 2, \\ Cn^{-2}(T^{-1}N^r)^{\alpha-\gamma} \ln n & \text{if } r(\gamma+1) = 2, \\ Cn^{-2-r(\alpha-\gamma)}(T^{-1}N^r)^{\alpha-\gamma} & \text{if } r(\gamma+1) > 2, \end{cases}$$

inequality (5.10) becomes

$$(5.18) \quad \sum_{k=\lceil n/2 \rceil}^{n-2} |T_{nk}| \leq Cn^{-(2-\alpha)-r(\alpha-\gamma)}(T^{-1}N^r)^{\alpha-\gamma},$$

inequality (5.13) becomes

$$(5.19) \quad |T_{n0}| \leq Cn^{-r\alpha}(T^{-1}N^r)^{\alpha-\gamma} \quad \text{for } n \geq 1,$$

and inequality (5.14) becomes

$$(5.20) \quad |T_{n,n-1}| \leq Cn^{-(2-\alpha)-r(\alpha-\gamma)}(T^{-1}N^r)^{\alpha-\gamma}.$$

In the last paragraph of the proof of Lemma 5.2 one replaces α by γ , and the conclusion of the lemma is now

$$|D_N^\alpha u(x_m, t_n) - D_t^\alpha u(x_m, t_n)| \leq CT^{\gamma-\alpha} N^{r(\alpha-\gamma)} n^{-\min\{2-\alpha+r(\alpha-\gamma), r\alpha\}}.$$

Consequently the convergence result of Theorem 5.3 becomes

$$(5.21) \quad \max_{(x_m, t_n) \in \bar{Q}} |u(x_m, t_n) - u_m^n| \leq C \left(T^\alpha h^2 + T^\gamma N^{-\min\{2-\alpha, r\gamma\}} \right)$$

for some constant $C = C(\gamma, \alpha)$.

In Remark 5.5 we do not consider the alternative situation where $\gamma > \alpha$, because—as shown in Remark 2.3—that hypothesis would place a severe restriction on the class of problems being studied.

Remark 5.6 (choice of mesh grading r). Numerical results in section 6 will show that Theorem 5.3 is sharp. It tells us that the optimal graded mesh is obtained when $r = (2 - \alpha)/\alpha$, as this then gives the highest possible rate of convergence

$O(h^2 + N^{-(2-\alpha)})$, while any larger value of r increases the temporal mesh width near $t = T$ and consequently increases the constant multiplier C . Furthermore, if the solution u of (1.1) satisfies the a priori bound (5.16) but the exact value of γ is unknown—let’s say one knows only that $\gamma \in [\alpha/2, \alpha]$ —then (5.21) indicates that it’s better to choose r too large than too small, so one should set $r = (2 - \alpha)/(\min \gamma) = 2(2 - \alpha)/\alpha$.

The discrete solution $\{u_m^n\}$ on the graded mesh can be extended globally by the bilinear interpolant

$$(5.22) \quad \bar{U}^{M,N}(x, t) := \sum_{m,n=0}^{M,N} u_m^n \phi_m(x) \psi_n(t) \quad \text{for } (x, t) \in \bar{Q},$$

where the $\phi_m(x)$ and $\psi_n(t)$ are piecewise linear basis functions with nodal values $\phi_k(x_i) = \delta_{i,k} = \psi_i(t_k)$.

COROLLARY 5.7. *Assume that $r\alpha \leq 2$. Then the global approximation $\bar{U}^{M,N}$ generated by the scheme (3.2) and the bilinear interpolant (5.22) satisfy*

$$(5.23) \quad \|u - \bar{U}^{M,N}\|_{L^\infty(\bar{Q})} \leq C \left(h^2 + T^\alpha N^{-\min\{2-\alpha, r\alpha\}} \right)$$

for some constant C .

Proof. Let \bar{u} be the bilinear interpolant to u on our graded mesh. On each mesh rectangle $R_{m,n} := (x_{m-1}, x_m) \times (t_{n-1}, t_n)$ one can easily derive the bound

$$\begin{aligned} \|u - \bar{u}\|_{L^\infty(R_{m,n})} &\leq Ch^2 \|u_{xx}\|_{L^\infty(R_{m,n})} \\ &\quad + C \min \left\{ \tau_n^2 \|u_{tt}\|_{L^\infty(R_{m,n})}, \max_{x \in [x_{m-1}, x_m]} \int_{s=t_{n-1}}^{t_n} |u_t(x, s)| ds \right\} \\ &\leq Ch^2 + C \min \{ \tau_n^2 t_{n-1}^{\alpha-2}, t_n^\alpha - t_{n-1}^\alpha \} \end{aligned}$$

by (2.9). For $n = 1$, this gives (since $t_1 = TN^{-r}$)

$$\|u - \bar{u}\|_{L^\infty(R_{m,1})} \leq C(h^2 + T^\alpha N^{-r\alpha});$$

and for $n > 1$ it gives, using (5.1) and the definition of t_{n-1} ,

$$\begin{aligned} \|u - \bar{u}\|_{L^\infty(R_{m,n})} &\leq Ch^2 + CT^\alpha N^{-r\alpha} n^{2(r-1)} (n-1)^{r(\alpha-2)} \\ &\leq Ch^2 + CT^\alpha N^{-r\alpha} (n-1)^{r\alpha-2} \\ &\leq Ch^2 + CT^\alpha N^{-r\alpha} \quad \text{because } r\alpha \leq 2. \end{aligned}$$

Hence $\|u - \bar{u}\|_{L^\infty(\bar{Q})} \leq C(h^2 + T^\alpha N^{-r\alpha})$.

The argument is completed by invoking the nodal error bound of Theorem 5.3 and a triangle inequality. \square

The assumption $r\alpha \leq 2$ of Corollary 5.7 is satisfied on our optimal mesh where $r = (2 - \alpha)/\alpha$.

6. Numerical experiments. In this section numerical results are presented for our scheme applied to two test problems of the form (1.1); in the first the exact solution is known, while the second is more complicated with an unknown exact solution.

Example 6.1.

$$(6.1) \quad D_t^\alpha u - \frac{\partial^2 u}{\partial x^2} = f(x, t) \quad \text{for } (x, t) \in (0, \pi) \times (0, 1]$$

with initial condition $u(x, 0) = 0$ for $0 < x < \pi$ and boundary conditions $u(0, t) = u(\pi, t) = 0$ for $0 \leq t \leq 1$. The function f in (6.1) is chosen such that the exact solution of the problem is $u(x, t) = (t^\alpha + t^3) \sin x$. This solution displays typical layer behavior at $t = 0$, i.e., its derivatives agree exactly with the bounds (2.4).

In Tables 3, 4, and 5 we use the format described in section 3.1 and show the maximum errors and the orders of convergence in the computed solutions for Example 6.1 with $\alpha = 0.8$, $\alpha = 0.6$, and $\alpha = 0.4$, respectively, and several values of the mesh grading exponent r .

These numerical results agree precisely with the theoretical rate of convergence of Theorem 5.3 for various values of r : in the case of a uniform mesh ($r = 1$) we observe

TABLE 3
Maximum errors and orders of convergence for Example 6.1 with $\alpha = 0.8$.

	N=M=64	N=M=128	N=M=256	N=M=512	N=M=1024	N=M=2048	N=M=4096
$r = 1$	5.335E-3 0.754	3.164E-3 0.770	1.856E-3 0.783	1.079E-3 0.790	6.240E-4 0.793	3.602E-4 0.796	2.075E-4
$r = \frac{(2-\alpha)}{2\alpha}$	1.100E-2 0.556	7.483E-3 0.574	5.028E-3 0.583	3.356E-3 0.585	2.237E-3 0.589	1.487E-3 0.593	9.863E-4
$r = \frac{2-\alpha}{\alpha}$	7.984E-3 1.201	3.473E-3 1.202	1.510E-3 1.202	6.563E-4 1.202	2.853E-4 1.202	1.240E-4 1.202	5.391E-5
$r = \frac{2(2-\alpha)}{\alpha}$	1.696E-2 1.177	7.501E-3 1.186	3.296E-3 1.192	1.443E-3 1.195	6.300E-4 1.197	2.747E-4 1.198	1.197E-4

TABLE 4
Maximum errors and orders of convergence for Example 6.1 with $\alpha = 0.6$.

	N=M=64	N=M=128	N=M=256	N=M=512	N=M=1024	N=M=2048	N=M=4096
$r = 1$	1.592E-2 0.566	1.075E-2 0.577	7.209E-3 0.585	4.807E-3 0.590	3.193E-3 0.593	2.117E-3 0.596	1.401E-3
$r = \frac{(2-\alpha)}{2\alpha}$	1.075E-2 0.674	6.740E-3 0.684	4.195E-3 0.690	2.601E-3 0.694	1.608E-3 0.696	9.923E-4 0.698	6.118E-4
$r = \frac{2-\alpha}{\alpha}$	5.119E-3 1.371	1.980E-3 1.381	7.601E-4 1.388	2.904E-4 1.392	1.106E-4 1.395	4.207E-5 1.397	1.598E-5
$r = \frac{2(2-\alpha)}{\alpha}$	1.225E-2 1.331	4.867E-3 1.356	1.901E-3 1.372	7.343E-4 1.382	2.817E-4 1.389	1.076E-4 1.393	4.098E-5

TABLE 5
Maximum errors and orders of convergence for Example 6.1 with $\alpha = 0.4$.

	N=M=64	N=M=128	N=M=256	N=M=512	N=M=1024	N=M=2048	N=M=4096
$r = 1$	3.357E-2 0.348	2.637E-2 0.360	2.055E-2 0.369	1.592E-2 0.376	1.227E-2 0.381	9.417E-3 0.386	7.207E-3
$r = \frac{(2-\alpha)}{2\alpha}$	7.207E-3 0.781	4.195E-3 0.789	2.428E-3 0.794	1.401E-3 0.796	8.066E-4 0.798	4.640E-4 0.799	2.667E-4
$r = \frac{2-\alpha}{\alpha}$	4.145E-3 1.511	1.455E-3 1.537	5.011E-4 1.555	1.705E-4 1.568	5.752E-5 1.576	1.929E-5 1.582	6.441E-6
$r = \frac{2(2-\alpha)}{\alpha}$	1.072E-2 1.444	3.940E-3 1.494	1.399E-3 1.526	4.855E-4 1.547	1.661E-4 1.561	5.631E-5 1.569	1.898E-5

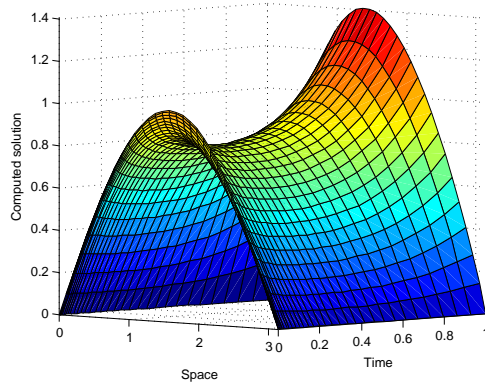


FIG. 2. Computed solution $u(x, t)$ of Example 6.2 when $\alpha = 0.6$.

a rate of α , while for a grading of $r = (2 - \alpha)/(2\alpha)$ we observe a rate of $(2 - \alpha)/2$, and when $r \geq (2 - \alpha)/\alpha$, the optimal rate of $2 - \alpha$ is obtained. Thus Theorem 5.3 is sharp.

In particular, Theorem 5.3 implies that one should choose $r \geq (2 - \alpha)/\alpha$ to obtain the best possible order of convergence $O(h^2 + N^{-(2-\alpha)})$, and our results confirm this.

Note that the maximum errors are greater for the choice of $r = 2(2 - \alpha)/\alpha$ than for $r = (2 - \alpha)/\alpha$. While the optimal rate of convergence $2 - \alpha$ is observed for both these choices, the larger value of r will concentrate more mesh points near $t = 0$, so the mesh is coarser away from $t = 0$, giving larger errors there.

Example 6.2. Consider the test problem

$$(6.2a) \quad D_t^\alpha u - \frac{\partial^2 u}{\partial x^2} + (1 + x)u = x(\pi - x)(1 + t^4) + t^2$$

for $(x, t) \in Q := (0, \pi) \times (0, 1]$ with

$$(6.2b) \quad u(0, t) = 0, \quad u(\pi, t) = 0 \quad \text{for } t \in (0, 1],$$

$$(6.2c) \quad u(x, 0) = \sin x \quad \text{for } x \in [0, \pi].$$

In Figure 2 the computed solution for $\alpha = 0.6$ is displayed. There is a layer at $t = 0$ and numerical experiments show that this layer becomes sharper as α decreases.

The exact solution of Example 6.2 is unknown, so the orders of convergence in the computed solutions are estimated using the two-mesh principle [6], which we now describe. Let u_m^n with $0 \leq m \leq M$ and $0 \leq n \leq N$ be the solution computed by our scheme (3.2) on the mesh defined at the beginning of section 3. Now consider a second mesh that is uniform in the spatial direction with $2M$ mesh intervals, and in the temporal direction is defined by

$$t_n = (n/(2N))^r \quad \text{for } 0 \leq n \leq 2N.$$

The solution computed by our scheme on this mesh is denoted by z_m^n with $0 \leq m \leq 2M$ and $0 \leq n \leq 2N$. Now the maximum two-mesh differences are defined by

$$D^{M,N} := \max_{0 \leq m \leq M} \max_{0 \leq n \leq N} |u_m^n - z_{2m}^{2n}|,$$

TABLE 6

Maximum two-mesh differences and orders of convergence for Example 6.2 with $\alpha = 0.8$.

	N=M=64	N=M=128	N=M=256	N=M=512	N=M=1024
$r = 1$	3.807E-003 0.772	2.229E-003 0.801	1.280E-003 0.809	7.303E-004 0.811	4.162E-004
$r = \frac{(2-\alpha)}{2\alpha}$	7.230E-003 0.600	4.769E-003 0.595	3.158E-003 0.625	2.048E-003 0.635	1.319E-003
$r = \frac{2-\alpha}{\alpha}$	2.561E-003 1.179	1.131E-003 1.188	4.966E-004 1.193	2.173E-004 1.196	9.485E-005
$r = \frac{2(2-\alpha)}{\alpha}$	5.284E-003 1.152	2.378E-003 1.173	1.055E-003 1.184	4.641E-004 1.191	2.033E-004

TABLE 7

Maximum two-mesh differences and orders of convergence for Example 6.2 with $\alpha = 0.6$.

	N=M=64	N=M=128	N=M=256	N=M=512	N=M=1024
$r = 1$	9.052E-003 0.498	6.410E-003 0.538	4.414E-003 0.564	2.986E-003 0.579	1.999E-003
$r = \frac{(2-\alpha)}{2\alpha}$	6.655E-003 0.628	4.307E-003 0.660	2.725E-003 0.678	1.703E-003 0.688	1.057E-003
$r = \frac{2-\alpha}{\alpha}$	1.808E-003 1.342	7.133E-004 1.363	2.773E-004 1.377	1.068E-004 1.385	4.087E-005
$r = \frac{2(2-\alpha)}{\alpha}$	4.235E-003 1.290	1.732E-003 1.332	6.877E-004 1.358	2.683E-004 1.373	1.036E-004

TABLE 8

Maximum two-mesh differences and orders of convergence for Example 6.2 with $\alpha = 0.4$.

	N=M=64	N=M=128	N=M=256	N=M=512	N=M=1024
$r = 1$	1.351E-002 0.172	1.200E-002 0.213	1.035E-002 0.249	8.711E-003 0.281	7.171E-003
$r = \frac{(2-\alpha)}{2\alpha}$	5.167E-003 0.697	3.186E-003 0.739	1.909E-003 0.764	1.124E-003 0.779	6.547E-004
$r = \frac{2-\alpha}{\alpha}$	1.497E-003 1.466	5.417E-004 1.509	1.904E-004 1.536	6.563E-005 1.555	2.234E-005
$r = \frac{2(2-\alpha)}{\alpha}$	3.789E-003 1.381	1.455E-003 1.457	5.300E-004 1.503	1.870E-004 1.532	6.466E-005

and they are used to compute the estimated rate of convergence

$$\log_2 \left(\frac{D^{M,N}}{D^{2M,2N}} \right).$$

The maximum two-mesh differences and their corresponding estimated orders of convergence for Example 6.2 with $\alpha = 0.8, 0.6, 0.4$ are given in Tables 6–8. The numerical results are in agreement with Theorem 5.3: they show that $r = (2 - \alpha)/\alpha$ yields the optimal rate of convergence $O(N^{-(2-\alpha)})$ together with the smallest maximum two-mesh differences.

REFERENCES

- [1] H. BRUNNER, *Collocation methods for Volterra Integral and Related Functional Differential Equations*, Cambridge Monogr. Appl. Comput. Maths. 15, Cambridge University Press, Cambridge, 2004, <https://doi.org/10.1017/CBO9780511543234>.

- [2] H. BRUNNER, L. LING, AND M. YAMAMOTO, *Numerical simulations of 2D fractional subdiffusion problems*, J. Comput. Phys., 229 (2010), pp. 6613–6622, <https://doi.org/10.1016/j.jcp.2010.05.015>.
- [3] R. COURANT AND D. HILBERT, *Methods of Mathematical Physics. Vol. I*, Interscience Publishers, New York, 1953.
- [4] E. CUESTA, C. LUBICH, AND C. PALENCIA, *Convolution quadrature time discretization of fractional diffusion-wave equations*, Math. Comp., 75 (2006), pp. 673–696, <https://doi.org/10.1090/S0025-5718-06-01788-1>.
- [5] K. DIETHELM, *The Analysis of Fractional Differential Equations*, Lecture Notes in Math., Springer, Berlin, 2010.
- [6] P. A. FARRELL, A. F. HEGARTY, J. J. H. MILLER, E. O’RIORDAN, AND G. I. SHISHKIN, *Robust Computational Techniques for Boundary Layers*, Appl. Math. 16, Chapman & Hall/CRC, Boca Raton, FL, 2000.
- [7] D. HENRY, *Geometric Theory of Semilinear Parabolic Equations*, Lecture Notes in Math. 840, Springer, Berlin, 1981.
- [8] B. JIN, R. LAZAROV, J. PASCIAK, AND Z. ZHOU, *Error analysis of semidiscrete finite element methods for inhomogeneous time-fractional diffusion*, IMA J. Numer. Anal., 35 (2015), pp. 561–582, <https://doi.org/10.1093/imanum/dru018>.
- [9] B. JIN, R. LAZAROV, AND Z. ZHOU, *An analysis of the L1 scheme for the subdiffusion equation with nonsmooth data*, IMA J. Numer. Anal., 36 (2016), pp. 197–221, <https://doi.org/10.1093/imanum/dru063>.
- [10] B. JIN, R. LAZAROV, AND Z. ZHOU, *Two fully discrete schemes for fractional diffusion and diffusion-wave equations with nonsmooth data*, SIAM J. Sci. Comput., 38 (2016), pp. A146–A170, <https://doi.org/10.1137/140979563>.
- [11] B. JIN AND Z. ZHOU, *An analysis of Galerkin proper orthogonal decomposition for subdiffusion*, ESAIM Math. Model. Numer. Anal., 51 (2017), pp. 89–113, <https://doi.org/10.1051/m2an/2016017>.
- [12] X. LI AND C. XU, *Existence and uniqueness of the weak solution of the space-time fractional diffusion equation and a spectral method approximation*, Commun. Comput. Phys., 8 (2010), pp. 1016–1051, <https://doi.org/10.4208/cicp.020709.221209a>.
- [13] Y. LIN AND C. XU, *Finite difference/spectral approximations for the time-fractional diffusion equation*, J. Comput. Phys., 225 (2007), pp. 1533–1552, <https://doi.org/10.1016/j.jcp.2007.02.001>.
- [14] F. LIU, P. ZHUANG, AND K. BURRAGE, *Numerical methods and analysis for a class of fractional advection-dispersion models*, Comput. Math. Appl., 64 (2012), pp. 2990–3007, <https://doi.org/10.1016/j.camwa.2012.01.020>.
- [15] Y. LUCHKO, *Maximum principle for the generalized time-fractional diffusion equation*, J. Math. Anal. Appl., 351 (2009), pp. 218–223, <https://doi.org/10.1016/j.jmaa.2008.10.018>.
- [16] Y. LUCHKO, *Initial-boundary-value problems for the one-dimensional time-fractional diffusion equation*, Fract. Calc. Appl. Anal., 15 (2012), pp. 141–160, <https://doi.org/10.2478/s13540-012-0010-7>.
- [17] R. METZLER AND J. KLAFTER, *The random walk’s guide to anomalous diffusion: a fractional dynamics approach*, Phys. Rep., 339 (2000), pp. 1–77, [https://doi.org/10.1016/S0370-1573\(00\)00070-3](https://doi.org/10.1016/S0370-1573(00)00070-3).
- [18] I. PODLUBNY, *Fractional Differential Equations*, Math. Sci. Eng., Academic Press, San Diego, CA, 1999.
- [19] K. SAKAMOTO AND M. YAMAMOTO, *Initial value/boundary value problems for fractional diffusion-wave equations and applications to some inverse problems*, J. Math. Anal. Appl., 382 (2011), pp. 426–447, <https://doi.org/10.1016/j.jmaa.2011.04.058>.
- [20] M. STYNES, *Too much regularity may force too much uniqueness*, Fract. Calc. Appl. Anal., 19 (2016), pp. 1554–1562.