Stability and non-negativity preservation for the numerical solution of space-fractional diffusion

Menghis T. Bahlibi

Supervisors: Ferenc Izsák, János Karátson

Institute of Mathematics
Faculty of Science, Eötvös Loránd University

Outline of the talk

Model problem

Space-fractional diffusion problem:

$$\begin{cases} \partial_t u(t,\mathbf{x}) = \mu \cdot -(-\Delta)^\alpha u(t,\mathbf{x}), & \mathbf{x} \in \Omega,\ t > 0 \\ u(0,\mathbf{x}) = u_0(\mathbf{x}), & \mathbf{x} \in \Omega, \end{cases} \tag{1}$$

where

- $\Omega \subset \mathbf{R}^d$ denotes the computational domain and $\mu \in \mathbf{R}^+$ is a diffusion coefficient.
- $(-\Delta)^\alpha$ denotes the fractional power of the negative Dirichlet Laplace operator.

For any $\alpha \in \mathbf{R}^+$ and $u \in C_0^\infty(\Omega)$, the operator $(-\Delta)^\alpha$ is defined as follows:

$$(-\Delta)^\alpha u(t, \mathbf{x}) = \sum_{k=1}^{\infty} u_k \lambda_k^\alpha \phi_k, \tag{2}$$

where

- $\{\phi_k\}_{k\in\mathbb{N}}$ is the orthonormal system of eigenfunctions of $-\Delta$,
- $\{\lambda_k\}_{k\in\mathbb{N}}$ corresponding eigenvalues,
- $u_k = \int_\Omega u\phi_k,\ k \in \mathbb{N}$ denotes the related Fourier coefficients.

For the details, we refer to [2], [5], and [6].

Finite difference discretization

To approximate $(-\Delta)^{\alpha}$, we apply FDM:

- Let $A \in \mathbf{R}^{n \times n}$ is any discretized form of $-\Delta$.

  The basic idea of the matrix transformation method is that

- for $A \approx -\Delta$, we also have $A^{\alpha} \approx (-\Delta)^{\alpha}$.

Using the matrix transformation method, (1) can be semi-discretized as

$$\partial_t \underline{u}(t, \cdot) = \mu \cdot -(A)^{\alpha} \underline{u}(t, \cdot). \tag{3}$$

Full discretization (θ- scheme)

We investigate the weighted average of the explicit and implicit finite difference schemes, leading to the so-called θ schemes:

$$\frac{\mathbf{u}^{n+1} - \mathbf{u}^n}{\delta} = -(A)^\alpha \mathbf{u}^n \cdot (1-\theta) + -(A)^\alpha \mathbf{u}^{n+1} \cdot \theta, \tag{4}$$

where

- δ is a time step and $\theta \in [0, 1]$ is a parameter.
- For $\theta = 0$, we get an explicit scheme, otherwise an implicit one.
- The choice $\theta = \frac{1}{2}$ leads to the Crank–Nicolson scheme.

Theoretical analysis

Definition

The matrix $B \in \mathbf{R}^{n \times n}$ is said to be an M matrix if

- all its off-diagonal elements are non-positive,
- there exists a vector $\mathbf{g} \in \mathbf{R}^n > 0$ such that $B\mathbf{g} > 0$ (element-wise positive).

Definition

The matrix $B \in \mathbf{R}^{n \times n}$ is called a Stieltjes matrix if

- B is a real symmetric positive definite matrix,
- all its off-diagonal elements are non-positive.

Note that a matrix $B \in \mathbf{R}^{n \times n}$ is called (weakly) diagonally dominant if

$$\sum_{k \neq j}^{n} |B_{jk}| \leq |B_{jj}| \quad \text{for all } j = 1, \ldots, n.$$

Theorem regarding Stieltjes matrix

We use the following key properties in the forthcoming analysis.

Theorem

Let B be a Stieltjes matrix and $\alpha \in (0, 1]$. In this case,

- *B is an M matrix such that it has a non-negative inverse,*
- *B^{α} is also a Stieltjes matrix.*

For the details, we refer to [4].

General assumptions

(i) A is symmetric positive definite.

(ii) A has positive diagonal elements and non-positive off-diagonals elements.

(iii) A is (weakly) diagonally dominant.

(iv) The diagonal of A is $d \cdot I$ with $d \geq 1$, where I is the identity matrix.

(v) The inequality $1 - \delta \cdot d \geq 0$ is satisfied for the time step δ in (5).

The explicit Euler discretization i.e. the case of $\theta = 0$ in (4), is given by

$$\mathbf{u}^{n+1} = \mathbf{u}^n - \delta A^\alpha \mathbf{u}^n = (I - \delta A^\alpha)\mathbf{u}^n. \tag{5}$$

Explicit Time stepping

Theorem

Assume that conditions (i)-(v) are satisfied for the time step δ in (5). Then the explicit time stepping in (5) is stable and preserves non-negativity for any $\alpha \in [0, 1]$.

Non-negativity for explicit time stepping

Theorem

Assume that for $\alpha = 1$, we have the condition $\delta \leq c_* h^2$ with some $c_* \in \mathbf{R}^+$ for preserving non-negativity in (5). Then for any $\alpha \in [0, 1]$, the condition $\delta \leq c_* h^{2\alpha}$ in (5) is sufficient for the same purpose.

Remarks for the Theorems

- In case of the conventional d-dimensional finite difference discretization, for $\alpha = 1$, we have the condition $\delta \leq \frac{h^2}{2d}$ for preserving non-negativity, see Theorem 6.2 in the article by István Faragó and Róbert Horváth [1].

- According to Theorem 3, we obtain the condition $\delta \leq \frac{h^{2\alpha}}{2d}$ for a general $\alpha \in [0, 1]$ using the matrix transformation technique.

- Taking a small discretization parameter h and $\alpha < 1$, this can be again a weaker condition for δ compared to the case of conventional diffusion.

Implicit Time stepping

Implicit Euler discretization:

$$\mathbf{u}^{n+1} = \mathbf{u}^n - \delta A^\alpha \mathbf{u}^{n+1}. \tag{6}$$

Theorem

Assume that the conditions (i) and (ii) are satisfied, then implicit time stepping is stable for any $\alpha \in [0,1]$ and preserves non-negativity.

- From (6), we have

$$\mathbf{u}^{n+1} = [I + \delta \cdot A^\alpha]^{-1}\mathbf{u}^n, \tag{7}$$

  where A^α is a Stieltjes matrix by the Theorem.

- Here $[I + \delta \cdot A^\alpha]^{-1}$ is element-wise positive.
- This results nonnegativity preservation.

Crank–Nicolson scheme

We also investigate the Crank–Nicolson scheme

$$\mathbf{u}^{n+1} = \mathbf{u}^n + \frac{\delta}{2} A^\alpha (\mathbf{u}^n + \mathbf{u}^{n+1}), \tag{8}$$

which can be rephrased as

$$\mathbf{u}^{n+1} = \left(I + \frac{\delta}{2} A^\alpha \right)^{-1} \left(I - \frac{\delta}{2} A^\alpha \right) \mathbf{u}^n. \tag{9}$$

Theorem

The Crank–Nicolson scheme in (9) is unconditionally stable and preserves non-negativity for $\delta < 2C_* \cdot h^{2\alpha} \quad \forall \alpha \in [0,1]$ *and for some* $C_* \in \mathbf{R}^+$.

Computation of matrix-vector products

- Matrix-vector products: $A^{\alpha}\mathbf{v}$, where $A \in \mathbf{R}^{n \times n}$ is a symmetric positive definite sparse matrix, $\mathbf{v} \in \mathbf{R}^{n}$ is an arbitrary vector and $\alpha \in [0, 1]$.
- We apply the Matlab codes described in the article by Ferenc Izsák and Béla J. Szekeres [3].
- Using this technique, we avoid the calculation of the matrix power A^{α}, which is a large and full matrix.

In the framework of the above technique, we should compute only sparse matrix-vector products, which leads to an efficient implementation.

Computation of matrix-vector products

- Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are the eigenvectors and $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ the corresponding eigenvalues of A.
- Then A^α has the same eigenvectors, while the corresponding eigenvalues are $\lambda_1^\alpha \leq \lambda_2^\alpha \leq \ldots \leq \lambda_n^\alpha$.
- For any vector $\mathbf{v} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \ldots + a_n\mathbf{x}_n \in \mathbf{R}^n$.
- The matrix-vector product is given as follows:

$$A^\alpha \mathbf{v} = a_1\lambda_1^\alpha \mathbf{x}_1 + a_2\lambda_2^\alpha \mathbf{x}_2 + \ldots + a_n\lambda_n^\alpha \mathbf{x}_n, \tag{10}$$

  which has to be approximated.
- To reduce the computational complexity, we compute only a few eigenvectors: $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{k_1}$ and $\mathbf{x}_{n-k_2+1}, \mathbf{x}_{n-k_2+2}, \ldots, \mathbf{x}_n$ along with the corresponding eigenvalues.

Truncated Taylor's approximation

This can be performed rather quickly using sparse eigensolvers of Matlab and can be used in the course of the entire simulation process.

The approximation has then two main steps:

- We first compute the component

$$\mathbf{v}_1 = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_{k_1}\mathbf{x}_{k_1} + a_{n-k_2+1}\mathbf{x}_{n-k_2+1} + \cdots + a_n\mathbf{x}_n.$$

  For this part, we can easily apply (10) to get

$$A^\alpha\mathbf{v}_1 = a_1\lambda_1^\alpha\mathbf{x}_1 + \cdots + a_{k_1}\lambda_{k_1}^\alpha\mathbf{x}_{k_1} + a_{n-k_2+1}\lambda_{n-k_2+1}^\alpha\mathbf{x}_{n-k_2+1} + \cdots + a_n\lambda_n^\alpha\mathbf{x}_n.$$

- For the rest $\mathbf{v}_2 = \mathbf{v} - \mathbf{v}_1$, we apply a truncated Taylor's approximation

$$A^\alpha\mathbf{v}_2 \approx \left(\frac{\sigma(A)}{2}\right)^\alpha \sum_{k=1}^{K} \binom{\alpha}{k} \left(\frac{2A}{\sigma(A)} - I\right)^k \mathbf{v}_2,$$

where σ denotes the spectral radius such that $\sigma\left(\frac{2A}{\sigma(A)} - I\right) \leq 1$ and the related binomial series is convergent.

Finally, we obtain the approximation $A^\alpha\mathbf{v} \approx A^\alpha\mathbf{v}_1 + A^\alpha\mathbf{v}_2$.

Computational results

The computations in the forthcoming examples will confirm our theoretical expectations in Theorem 2 and Theorem 3.

Example 1:1D space-fractional diffusion problem:

$$\begin{cases} \partial_t u(t,x) = -(-\Delta)^\alpha u(t,x), & x \in (0,\pi/2), \quad t \in (0,0.1) \\ u(0,x) = \sin(2x), & x \in (0,\pi/2) \\ u(t,0) = u(t,\pi/2) = 0, & t \in (0,0.1), \end{cases} \tag{11}$$

- Analytic solution is given by $u(t,x) = e^{-4^\alpha t} \sin 2x$.
- $\alpha \in [0,1]$ denotes the exponent in the matrix power, k_1 and k_2 are the numbers of eigenvectors of the matrix that are used for the decomposition and K is the number of terms in Taylor's approximation.
- n is the number of subintervals in $\Omega = (0,\pi/2)$, δ is the time step, where the domain is discretized using a uniform grid with the grid size $h = \frac{\pi}{2n}$.

Experimental error rates with respect to L_2-norm

δ	h	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$	r_1	r_2	r_3
10^{-4}	0.2618	1.3×10^{-3}	2.4×10^{-3}	3.8×10^{-3}	-	-	-
10^{-4}	0.1309	4.143×10^{-4}	6.97×10^{-4}	1.1×10^{-3}	1.6	1.7	1.7
10^{-4}	0.0654	1.221×10^{-4}	2.158×10^{-4}	3.235×10^{-4}	1.7	1.6	1.7
10^{-4}	0.0327	3.817×10^{-5}	7.192×10^{-5}	8.987×10^{-5}	1.6	1.5	1.8
10^{-4}	0.0164	1.272×10^{-5}	2.258×10^{-5}	2.248×10^{-5}	1.5	1.6	1.9

- In this Convergence results, the first $K = 1000$ terms in the Taylor expansion and $k_1 = k_2 = 20$ were computed.
- Here r_1, r_2, and r_3 are the convergence order with respect to the space parameter h in case of $\alpha = 0.4$, $\alpha = 0.6$ and $\alpha = 0.8$, respectively.

Non-negativity preservation is the consequence of the stability condition $\delta \leq \frac{h^{2\alpha}}{2}$, since we prescribed non-negative initial and boundary conditions.

Stability condition

The stability conditions for the maximum time step δ_{exp} in the experiments and the maximal time step $\delta_{\text{theory}} = \frac{h^{2\alpha}}{2}$ ensuring stability according to Theorem 3 for the numerical solution of (11) are recorded in Table 2.

	$\alpha = 0.5$			$\alpha = 0.7$			$\alpha = 0.9$	
h	0.1	0.01	0.001	0.1	0.01	0.001	0.1	0.01
δ_{exp}	0.048	0.0047	0.00047	0.0195	0.0006	0.0000031	0.006	0.000123
δ_{theory}	0.05	0.005	0.0005	0.0199	0.0008	0.000032	0.008	0.000126

- Comparison of experimental stability results and the result of Theorem 3 for different values of h and fractional powers α.

Non-negativity preservation

Example 2: 2D space-fractional diffusion problem:

$$\begin{cases} \partial_t u(t,x) = -(-\Delta)^{\alpha} u(t,x,y), & (x,y) \in \Omega, \quad t \in (0, 0.1) \\ u(t,x,y) = 0, \quad (x,y) \in \partial\Omega, \quad t \in (0, 0.1) \\ u(0,x,y) = \sin\, x \sin\, y, \quad (x,y) \in \Omega, \end{cases} \tag{12}$$

where

- $\Omega = (0,\pi) \times (0,\pi)$ and the analytic solution is given by $u(t,x,y) = e^{-2^{\alpha} t}(\sin\, x \sin\, y)$.
- Here, we have used a uniform square grid of size $h = \frac{\pi}{n}$ for the finite difference discretization of Ω.
- The rest of the notations coincide with the ones in Example 1.

Experimental error rates with respect to L_2-norm

δ	h	$\alpha = 0.2$	$\alpha = 0.5$	$\alpha = 0.6$	$\alpha = 0.8$
0.02	0.5236	3.16×10^{-2}	3.87×10^{-2}	4.13×10^{-2}	4.7×10^{-2}
0.005	0.2618	7.8×10^{-3}	9.6×10^{-3}	1.02×10^{-2}	1.16×10^{-2}
0.00125	0.1309	1.9×10^{-3}	2.4×10^{-3}	2.5×10^{-3}	2.9×10^{-3}
0.0003125	0.0654	4.83×10^{-4}	5.9×10^{-4}	5.62×10^{-4}	7.26×10^{-4}
0.000078125	0.0327	1.23×10^{-4}	1.5×10^{-4}	1.61×10^{-4}	1.81×10^{-4}

- Convergence order is at least $O(\delta)+O(h^2)$, the first $K = 1000$ terms in the Taylor expansion, and $k_1 = k_2 = 20$ were computed.

Non-negativity preservation is the consequence of the stability condition $\delta \leq \frac{h^{2\alpha}}{4}$, since we prescribed non-negative initial and boundary conditions.

Stability condition + convergence order

Example 3: 2D space-fractional diffusion problem:

$$\begin{cases} \partial_t u(t,x) = -(-\Delta)^\alpha u(t,x,y), & (x,y) \in \Omega, \quad t \in (0, 0.1) \\ u(t,x,y) = 0, & (x,y) \in \partial\Omega, \quad t \in (0, 0.1) \\ u(0,x,y) = \sin x \sin 2y + \sin 2x \sin y, & (x,y) \in \Omega, \end{cases} \tag{13}$$

where

- $\Omega = (0,\pi) \times (0,\pi)$ and the analytic solution is given by $u(t,x,y) = e^{-5^\alpha t}(\sin x \sin 2y + \sin 2x \sin y)$.
- We use the same notations as in Example 2.

Experimental error rates with respect to L_2-norm

δ	h	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$	$\alpha = 0.9$
0.02	0.5236	7.49×10^{-2}	0.1001	0.1293	0.1445
0.005	0.2618	1.86×10^{-2}	2.48×10^{-2}	3.18×10^{-2}	3.54×10^{-2}
0.00125	0.1309	4.6×10^{-3}	6.2×10^{-3}	7.9×10^{-3}	8.8×10^{-3}
0.0003125	0.0654	1.1×10^{-3}	1.5×10^{-3}	2.0×10^{-3}	2.2×10^{-3}
0.000078125	0.03879	2.78×10^{-4}	3.81×10^{-4}	5.2×10^{-4}	5.56×10^{-4}

- Convergence order is at least $O(\delta)+O(h^2)$ and the first $K = 1000$ terms in the Taylor expansion and $k_1 = k_2 = 20$ were computed.

In the experiments, we have only stability $\delta \leq \frac{h^{2\alpha}}{4}$ since the given initial condition is negative at some points.

Stability condition

The stability conditions for the maximum time step $\delta_{\exp}$ in the experiments and the maximal time step $\delta_{\text{theory}} = \frac{h^{2\alpha}}{4}$ ensuring stability according to Theorem 3 for the numerical solution of (13) are recorded in Table 4.

	$\alpha = 0.5$			$\alpha = 0.7$		$\alpha = 0.9$	
h	0.1	0.01	0.001	0.1	0.01	0.1	0.01
$\delta_{\exp}$	0.0248	0.00245	0.000247	0.00993	0.000393	0.00392	0.000060
δ_{theory}	0.025	0.0025	0.00025	0.00995	0.000396	0.00396	0.000063

- Comparison of experimental stability results and the result of Theorem 3 for different values of h and fractional powers α.

Conclusion

In the framework of the matrix transformation method:

- The stability, and preservation of non-negativity for space-fractional diffusion problems can be analyzed independently from the spatial dimension.
- The stability condition depends only on the conventional finite difference discretization matrix and the fractional power of the Laplacian.
- Taking a spatial refinement, this leads to milder stability conditions compared to the case of conventional diffusion.
- The preservation of non-negativity can be established.

The corresponding manuscript was submitted.

References

[1] Istvan Farago and Robert Horvath. "Discrete maximum principle and adequate discretizations of linear parabolic problems". In: *SIAM Journal on Scientific Computing* 28.6 (2006), pp. 2313–2336.

[2] Stanislav Harizanov, Raytcho Lazarov, and Svetozar Margenov. "A survey on numerical methods for spectral space-fractional diffusion problems". In: *Fractional Calculus and Applied Analysis* 23.6 (2020).

[3] Ferenc Izsák and Béla J Szekeres. "Efficient computation of matrix power-vector products: Application for space-fractional diffusion problems". In: *Applied Mathematics Letters* 86 (2018), pp. 70–76.

[4] Charles A Micchelli and Ralph A Willoughby. "On functions which preserve the class of Stieltjes matrices". In: *Linear Algebra and its Applications* 23 (1979), pp. 141–156.

[5] Ricardo H Nochetto, Enrique Otárola, and Abner J Salgado. "A PDE approach to fractional diffusion in general domains: a priori error analysis". In: *Foundations of Computational Mathematics* 15 (2015), pp. 733–791.

[6] Béla J Szekeres and Ferenc Izsák. “Finite difference approximation of space-fractional diffusion problems: the matrix transformation method”. In: *Computers & Mathematics with Applications* 73.2 (2017), pp. 261–269.

Thank you for your attention!