

М. А. Новиков^{1✉}, В. В. Лисица¹

Численное решение параболического уравнения с использованием разложения матричной экспоненты по полиномам Чебышева в задачах моделирования физических процессов в многомасштабных средах

¹Институт математики им. С. Л. Соболева СО РАН, г. Новосибирск, Российская Федерация
e-mail: m.a.novikov@math.nsc.ru

Аннотация. Многие физические процессы описываются уравнениями параболического типа. К примеру, распространение тепла в сплошной среде и квазистатические процессы в рамках теории упругости и ее модификаций (включая пороупругость) описываются такими моделями. Стандартный конечно-разностный подход в соответствующих задачах во временной области зачастую требует малого шага по времени, что приводит к существенным вычислительным затратам. В данной статье предлагается альтернативный подход к численному решению параболических задач с использованием полиномов Чебышева. В частности, дифференциальное уравнение в частных производных параболического типа вначале пространственной дискретизацией сводится к системе обыкновенных дифференциальных уравнений по времени. В полученной задаче явное решение состоит из двух компонент, зависящих от матричной экспоненты, начальных условий и внешних источников. Численный метод основывается на приближении матричной экспоненты неполным разложением по матричным полиномам Чебышева с коэффициентами в форме модифицированных функций Бесселя первого рода. В работе приведено указанное приближенное решение исходного уравнения, кратко описан процесс его вычисления и предложен критерий остановки на основе необходимого числа первых слагаемых разложений обеих частей решения. На примере одномерной задачи теплопроводности выполнена верификация численного алгоритма сравнением с результатами использования классической явной конечно-разностной схемы. Кроме того, выполнено сравнение вычислительных затрат двух подходов на примере числа условных операций умножения матрицы на вектор.

Ключевые слова: теплопроводность, пороупругость, параболические уравнения, полиномы Чебышева

M. A. Novikov^{1✉}, V. V. Lisitsa¹

Numerical solution of the parabolic equation using Chebyshev polynomial decomposition for the matrix exponent for the modeling of physical phenomena in multi-scale media

¹Sobolev Institute of Mathematics SB RAS, Novosibirsk, Russian Federation
e-mail: m.a.novikov@math.nsc.ru

Abstract. Many physical phenomena are described by parabolic equations. Such cases include heat transport problems and also quasi-static processes within the theory of elasticity and its modifications including poroelasticity. Classical finite-difference approach in such problems usually requires quite small time step, and this results in significant computational cost. In this paper we present alternative approach to numerical solution of parabolic problems using Chebyshev polynomials. In particular, parabolic PDE can be reduced to the system of time dependence ODEs by a spatial discretization of

the solution. For the latter one explicit solution is formulated in terms of the matrix exponent, initial condition and source terms. Numerical method is based on the matrix exponent approximation by its incomplete Chebyshev series expansion with the coefficients in the form of modified Besseli fucttions of first type. We derive this solution approximation, describe its computation process and discuss the stopping criterion for it via bumbers of series first terms applied. The verification of the corresponding algorithm is performed for the one-dimensional heat conduction problem by the comparison of the approximate solutions by considered method and the standart explicit finite-difference scheme. Finally, numerical cost for both approaches is considered to demonstrate the potential economy in the number of required matrix-vector multiplication operations.

Keywords: heat conduction, poroelasticity, parabolic equations, Chebyshev polynomials

Введение

Параболические уравнения возникают во многих прикладных задачах, при математическом моделировании физических процессов в сплошных средах. Наиболее известны среди них задачи теплопроводности [1]. Однако уравнения теории упругости в квазистатической постановке также обладают параболической структурой [2, 3]. В обоих случаях современные задачи требуют как детального описания среды, так и больших по размеру расчетных областей. Поэтому необходимо использование вычислительных сеток с большим числом узлов при применении конечно-разностных методов [2, 3], конечно-элементных аппроксимаций [4, 5] и других. Зачастую условия устойчивости накладывают жёсткие ограничения на шаг по времени, что приводит к необходимости выполнения большого числа шагов [6, 7]. Таким образом, одной из ключевых задач при разработке современных численных алгоритмов является снижение вычислительных затрат при решении крупномасштабных задач.

В задачах вычислительной пороупругости в последние десятилетия существенную роль играет квазистатическая постановка соответствующих уравнений. В частности, исследование влияния структуры пористых флюидонасыщенных сред на сейсмические характеристики (такие как скорости распространения и затухание сейсмических волн) часто основывается на квазистатическом моделировании. Подобные задачи возникают, например, при сейсмическом мониторинге коллекторов углеводородов и анализе изменений свойств пород многолетней мерзлоты. Такие задачи входят в рамки численного апскейлинга, где отклик среды на внешнее нагружение усредняется для получения эффективных оценок фазовых скоростей и коэффициентов затухания [2-5]. При этом для решения уравнений в частотной области для каждой частоты расчёты необходимо выполнять отдельно. Временная же область позволяет проводить универсальные вычисления с получением частотных характеристик преобразованием Фурье.

Для численного апскейлинга применяются различные модели пороупругих сред. К числу наиболее известных относится феноменологическая модель Био [8, 9] и её обобщения на случай многофазных флюидов [10, 11] и вязкоупругого скелета [12]. Известны и более строгие математические модели на основе термодинамически согласованных систем законов сохранения [13, 14]. В динамической постановке эти модели представляют гиперболические системы, а при

переходе к квазистатической постановке [2,3] во временной области приводятся к параболическим уравнениям. В связи с этим разработка эффективных численных методов решения параболических задач важна для исследования зависимости затухания и распространения сейсмических волн от свойств разномасштабных флюидонасыщенных сред (пористость, трещиноватость, насыщение и т.д.).

В статье представлен подход к решению начально-краевой задачи для параболического уравнения, основанный на разложении матричной экспоненты по полиномам Чебышёва [15, 16]. Исходная задача для уравнения в частных производных преобразуется в систему обыкновенных дифференциальных уравнений, для которой известно аналитическое решение. Основная идея подхода заключается в представлении матричной экспоненты в решении в виде ряда матричных полиномов Чебышёва. Разложение матричной экспоненты по полиномам Чебышёва включает в качестве коэффициентов модифицированные функции Бесселя первого рода. Для численного решения вычисляется конечное число членов ряда, что дает приближение аналитического решения. Выполнена верификация подхода для частного случая постоянных граничных условий в исходной задаче для уравнения в частных производных при отсутствии внешних источников.

Методы и материалы

Рассматривается простейшая постановка одномерной задачи теплопроводности – начально-краевая задача в области $(t, x) \in (0, T] \times (0, L)$ для параболического уравнения с постоянными коэффициентами:

$$\begin{cases} \frac{\partial u(t, x)}{\partial t} - \frac{\partial^2 u(t, x)}{\partial x^2} = f(t, x), & x \in (0, L), \quad t \in (0, T]; \\ u(0, x) = \phi(x), & x \in [0, L]; \\ u(t, 0) = u_l(t), & t \in (0, T]; \\ u(t, L) = u_r(t), & t \in (0, T]. \end{cases} \quad (1)$$

Здесь $u(t, x)$ – искомая функция (температура), начальные условия заданы в $t = 0$ функцией $\phi(x)$, а краевые условия на концах одномерной области $[0, L]$ заданы функциями $u_l(t)$ и $u_r(t)$. При пространственной дискретизации расчетной области и соответствующей аппроксимации второй производной по пространству центральной разностью уравнение (1) сводится к системе обыкновенных дифференциальных уравнений вида:

$$\frac{\partial U(t)}{\partial t} + AU(t) = F(t), \quad t \in (0, T] \quad (2)$$

на вектор-функцию $U(t)$ значений поля $u(t, x)$ в узлах расчетной сетки, где матрица A при регулярной дискретизации постоянна, а вектор $F(t)$ в правой части образован узловыми значениями $f(t, x)$ и граничными условиями задачи (1). Известно, что решение системы (2) представимо в виде

$$U(t) = e^{-At}U_0 + \int_0^t e^{-A(t-s)}F(s) ds = e^{-At}U_0 + \int_0^t e^{-As}F(t-s) ds, \quad (3)$$

где вектор U_0 – вектор узловых значений функции начального условия $\phi(x)$ [17]. В обоих слагаемых решения присутствует матричная экспонента e^{-At} , при некотором приближении которой может быть приближено все решение. Поскольку зачастую (к примеру, в подзадачах численного апскейлинга [2,3]) граничные условия не зависят от времени, далее для упрощения рассматриваем $F(t) = F = const$. При этом вектор F выносится из интеграла во второе слагаемое.

Полиномы Чебышева обладают множеством важных свойств, обеспечивающих близкое к наилучшему приближение непрерывных функций (наименьшее отклонение от нуля на отрезке $[-1,1]$, наличие точек альтеранса и т.д.). С теоретическим обоснованием свойств полиномов Чебышева можно ознакомиться, к примеру, в книге [16]. Напомним, что полиномы Чебышева определяются рекурсивно следующим образом:

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_k(x) = 2xT_{k-1}(x) - T_{k-2}(x), \quad k = 2, 3, \dots \quad (4)$$

Формулы (4) могут быть распространены на матричную переменную. Существуют различные формы чебышевских полиномов. В частности, на отрезке $[-1,1]$ справедливо представление $T_k(x) = \cos(n \arccos x)$.

Разложение матричной экспоненты по полиномам Чебышева может быть выражено через функции Бесселя первого рода $J_k(z)$ с целыми индексами [16]. В частности, с использованием функции Якоби-Ангера и связи функций Бесселя $J_k(iz) = i^k I_k(z)$ с модифицированными функциями Бесселя первого рода I_k на интервале $[-1,1]$ может быть получено выражение для $x, q \in \mathbb{R}$:

$$e^{q \cos t} = 2 \sum_{k=0}^{\infty} ' I_k(q) \cos t \xrightarrow{t=\arccos x} e^{qt} = 2 \sum_{k=0}^{\infty} ' I_k(q) T_k(x). \quad (5)$$

Здесь $\sum'$ означает суммирование, где при первом слагаемом стоит множитель $1/2$. Поскольку матрица A в системе (2) симметрична, матричный аналог разложения (5) полностью определяется этой скалярной функцией на собственных значениях матрицы $x = \lambda_i \in \text{Spec}(A)$. Для использования формулы (5) они должны располагаться в интервале $[-1,1]$. Поэтому вначале необходимо нормировать спектр матрицы заменой

$$\tilde{A} = I - 2 \frac{A - \lambda_{\min} I}{\lambda_{\max} - \lambda_{\min}}. \quad (6)$$

Здесь $\lambda_{\min}$ и $\lambda_{\max}$ – наименьшее и наибольшее собственные значения матрицы A . С учетом этой замены справедливо разложение

$$e^{-At} = 2e^{-\lambda_+ t} \sum_{k=0}^{\infty} {}' I_k(\lambda_- t) T_k(\tilde{A}), \quad (7)$$

где λ_- и λ_+ – полуразность и полусумма наибольшего и наименьшего собственных значений A . Для получения приближения решения (3) экспоненты в первом слагаемом и подынтегральном выражении второго слагаемого заменяются частичными суммами разложения (7):

$$\tilde{U} = 2e^{-\lambda_+ t} \sum_{k=0}^{N_1} {}' I_k(\lambda_- t) T_k(\tilde{A}) U_0 + 2 \sum_{k=0}^{N_2} {}' \left[\int_0^t e^{-\lambda_+ s} I_k(\lambda_- s) ds \right] T_k(\tilde{A}) F. \quad (8)$$

Обозначим в $\tilde{U} = \tilde{U}_1 + \tilde{U}_2$. Основные вычисления приближенного решения (8) могут быть выражены в операциях умножения матрицы на вектор в процессе последовательного вычисления слагаемых обеих сумм. Из-за рекурсивной связи полиномов Чебышева для нахождения нового слагаемого суммы (7) необходимо лишь одно матричное умножение и несколько сложений, которые мы не включаем в грубую оценку вычислительной сложности. Так, для вычисления первого слагаемого решения (8) $\tilde{U}_1$ необходимо N_1 умножений матрица-вектор, для второго слагаемого $\tilde{U}_2 - N_2$ таких операций. Однако в $\tilde{U}_2$ также необходимо вычислить интегральные коэффициенты, что в зависимости от метода численного интегрирования может привносить различный вклад в вычислительную сложность.

В численном моделировании обычно выполняется расчет на множестве шагов по времени, поэтому приближенное решение (8) находится несколько раз для разных моментов времени $t_0, t_1, t_2, \dots, t_n$ с использованием решения на предыдущем моменте времени в качестве начального вектора U_0 . При постоянном шаге по времени Δt второе слагаемое $\tilde{U}_2$ в (8) находится один раз за всю процедуру, а $\tilde{U}_1$ считается на каждом шаге, поэтому окончательная оценка основных операций умножения матрица-вектор есть $n \cdot N_1 + N_2$.

Для оценки точности приближения решения рассматривается оценка абсолютной погрешности

$$\begin{aligned} \|U(t) - \tilde{U}\| &\leq 2e^{-\lambda_+ t} \sum_{k=N_1+1}^{\infty} {}' I_k(\lambda_- t) \|T_k(\tilde{A})\| \|U_0\| + \\ &+ 2 \sum_{k=N_2+1}^{\infty} {}' \left| \int_0^t e^{-\lambda_+ s} I_k(\lambda_- s) ds \right| \|T_k(\tilde{A})\| \|F\|. \end{aligned} \quad (9)$$

Отсюда при предположении об ограниченности $\|U_0\| \leq C_1$ и $\|F\| \leq C_2$ и оценке $\|T_k(x)\| \leq 1$ при $x \in [-1, 1]$, выражая полученные остатки рядов в слагаемых правой части (9) через матричную экспоненту $e^{\lambda_- t}$ и первые слагаемые рядов, выводятся условия для ограничения погрешности малым положительным ε для каждого слагаемого решения:

$$\begin{aligned} & \left(1 - 2e^{-\lambda_+ t} \sum_{k=0}^{N_1} I_k(\lambda_- t) \right) C_1 \leq \varepsilon, \\ & \left(t - 2 \sum_{k=0}^{N_2} \left[\int_0^t e^{-\lambda_+ s} I_k(\lambda_- s) ds \right] \right) C_2 \leq \varepsilon \end{aligned} \quad (10)$$

Условие (10) определяет необходимое число первых членов обоих рядов решения для достижения заданной точности.

Результаты

Для верификации представленного подхода с использованием соответствующего алгоритма получены решения одномерной задачи (1). В частности, получены решения в отсутствие источников $f(t, x) = 0$ для параметров $L = 1$, $u_l = u_r = 1$ при начальном условии, заданном квадратичной функцией $2 + (4(u_l - 2)/L^2) \cdot (x - L/2)^2$. Расчетная область представлена сеткой с шагом $h = 0.01$.

Для сравнения численного решения описанным методом задача (1) также решалась с помощью стандартной явной конечно-разностной схемы первого порядка по времени и второго по пространству. Для устойчивости схемы шаг по времени Δt выбран в соответствии условию Куранта [18]. Обоиими методами задача была решена для времени $T = n_T \Delta t$ – за n_T шагов явной схемой и за один шаг при помощи полиномов Чебышева. Рассмотрено три различных шага по времени $\Delta t = \{h^2/2, h^2/4, h^2/8\}$ для явной схемы, а в критерии (10) параметр $\varepsilon = 10^{-6}$ для сравнимой точности с точностью численного решения явной схемой. Относительная погрешность решений обоими методами в зависимости от n_T представлена на Рисунке 1. На Рисунке 2 показан рост числа операций умножения матрица-вектор с увеличением $T = n_T \cdot \Delta t$ для обоих методов и различных шагов по времени для предложенного подхода. В частности, для явной схемы можно считать $N_M = n_T$, а для предложенного подхода $N_M = N_1 + N_2$ из (8).

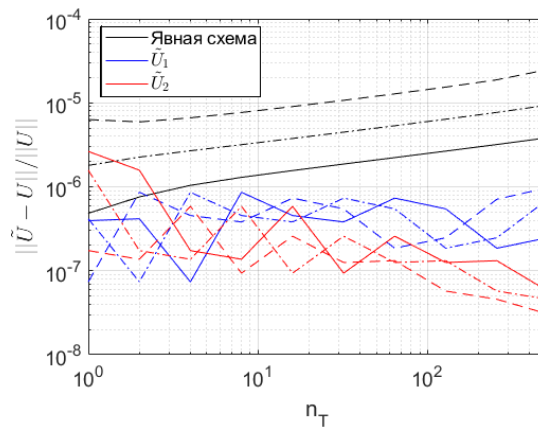


Рис. 1. Относительная погрешность численного решения задачи (1). Сплошная, штриховая и штрихпунктирная линии соответствуют шагам по времени явной схемы $h^2/2, h^2/4, h^2/8$.

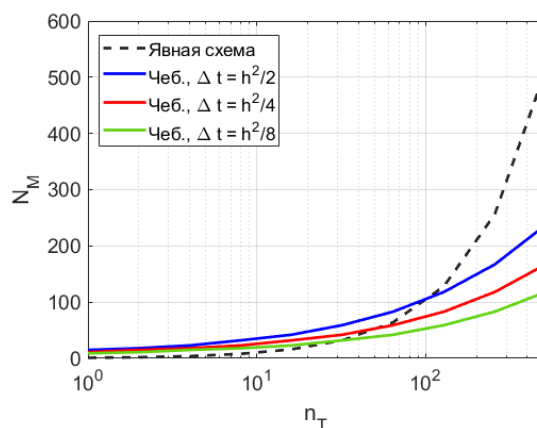


Рис. 2. Число основных операций для вычисления решения.

Обсуждение

Как видно на Рисунке 1, относительная погрешность предложенного метода в верификационных расчетах сравнима с погрешностью явной схемы. При этом рост числа основных операций, сопоставимых с умножением матрица-вектор в явной схеме имеет больший порядок, нежели рост для предложенного метода. Это приводит к тому, что несмотря на то, что для малого числа шагов по времени явная схема выгоднее, для всех трех рассматриваемых шагов по времени после превышения интервалом $[0, T]$ некоторого значения для вычислений предложенным подходом требуется меньше вычислений. Так, для $\Delta t = h^2/8$ число $N_1 + N_2 \approx 30$, для $\Delta t = h^2/4$ $N_1 + N_2 \approx 60$, а для $\Delta t = h^2/2$ $N_1 + N_2 \approx 100$. Поскольку для приближения обоих слагаемых в ходе расчетов требовалось приблизительно одинаковое число первых членов ряда, примерно 15, 30 и 50 первых полиномов Чебышева с неотрицательными индексами необходимо для сравнимой погрешности предложенного метода с явной схемой. Более того, для наибольшего числа шагов по времени явной схемы потребовалось приблизительно в два раза меньше основных операций умножения матрица-вектор. Однако, поскольку такая оценка числа операций достаточно груба, в дальнейшем необходим более детальный учет вычислительных затрат, к примеру, с учетом численного интегрирования для нахождения коэффициентов второго ряда решения (8).

Заключение

В работе представлен эффективный численный подход к решению начально-краевой задачи для уравнения параболического типа, к которым относятся как задачи теплопроводности, так и задачи квазистатического нагружения пороупругих образцов среды в рамках задачи апскейлинга. Подход основан на сведении задачи к системе обыкновенных дифференциальных уравнений на функции значений в узлах пространственной сетки с последующим приближением матричной экспоненты в решении частью ряда по полиномам Чебышева с коэффициентами в виде

модифицированных функций Бесселя первого рода. Представлен возможный критерий остановки вычислительного процесса для достижения заданной точности. Верификация численного метода на одномерной задаче сравнением со стандартным конечно-разностным подходом демонстрирует эффективность первого для относительно больших временных интервалов. В дальнейшем необходим более детальный анализ эффективности алгоритма. Также в дальнейшие планы входит применение представленного подхода к решению задач оценки затухания и дисперсии сейсмических волн в многофазных пористых средах.

Благодарности

Работа выполнена в рамках государственного задания Минобрнауки России для ИМ СО РАН (тема №FWNF-2026-0018).

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Бицадзе, А. В. Уравнения математической физики : учебное пособие. – 2-е изд., перераб. и доп. – М. : Наука, 1982. – 336 с.
2. Solovyev, S., Novikov, M., Lisitsa, V. A numerical investigation of wave-induced fluid flows in anisotropic fractured porous media // *Computers and Mathematics with Applications*. – 2023. – Vol. 140. – P. 78–88.
3. Novikov, M., Lisitsa, V., Romenski, E. Numerical Modeling of Fluid-Saturated Porous Sample Loading Based on Thermodynamically Compatible System Theory // *ICCSA 2024 Workshops / O. Gervasi et al. (Eds.)*. – Cham : Springer, 2024. – Vol. 14817. – P. 237–249. – DOI: 10.1007/978-3-031-65238-7_16.
4. Quintal, B., Steeb, H., Frehner, M., Schmalholz, S. Quasi-static finite element modeling of seismic attenuation and dispersion due to wave-induced fluid flow in poroelastic media // *Journal of Geophysical Research*. – 2011. – Vol. 116.
5. Rubino, J. G., Caspari, E., Muller, T. M., Milani, M., Barbosa, N. D., Holliger, K. Numerical upscaling in 2D heterogeneous poroelastic rocks: Anisotropic attenuation and dispersion of seismic waves // *Journal of Geophysical Research: Solid Earth*. – 2016. – Vol. 121, no. 9.
6. Lisitsa, V., Vishnevsky, D. Lebedev scheme for the numerical simulation of wave propagation in 3D anisotropic elasticity // *Geophysical Prospecting*. – 2010. – Vol. 58, no. 4. – P. 619–635.
7. Bernth, H., Chapman, C. Comparison of the dispersion relations for anisotropic elastodynamic finite-difference grids // *Geophysics*. – 2011. – Vol. 76, no. 3. – P. WA43–WA50.
8. Biot, M. A. Theory of propagation of elastic waves in fluid-saturated porous solid. I. low-frequency range // *J. Acoust. Soc. Am.* – 1956. – Vol. 28. – P. 168–178.
9. Biot, M. A. Theory of propagation of elastic waves in a fluid-saturated porous solid. II. higher frequency range // *J. Acoust. Soc. Am.* – 1956. – Vol. 28. – P. 179–191.
10. Santos, J. E., Douglas, J., Corbero, T., Lovera, O. M. A model for wave propagation in a porous medium saturated by a two-phase fluid // *J. Acoust. Soc. Am.* – 1990. – Vol. 87, no. 4. – P. 1439–1448.
11. Carcione, J. M., Cavallini, F., Santos, J. E., Ravazzoli, C. L., Gauzellino, P. M. Wave propagation in partially saturated porous media: simulation of a second slow wave // *Wave Motion*. – 2003. – Vol. 39. – P. 227–240.
12. Liu, X., Greenhalgh, S., Zhou, B., Greenhalgh, M. Effective biot theory and its generalization to poroviscoelastic models // *Geophysical Journal International*. – 2018. – Vol. 212, no. 2. – P. 1255–1273. – DOI: 10.1093/gji/ggx460.
13. Romenski E., Reshetova G., Peshkov I., Dumbser M. Modeling wavefields in saturated elastic porous media based on thermodynamically compatible system theory for two-phase solid-fluid mixtures // *Computers and Fluids*. – 2020. – Vol. 206. – P. 104587.

14. Romenski E., Reshetova G., Peshkov I., Dumbser M. Modeling wavefields in saturated elastic porous media based on thermodynamically compatible system theory for two-phase solid-fluid mixtures // *Computers and Fluids*. – 2020. – Vol. 206. – P. 104587.
15. Пашковский, С. Вычислительные применения многочленов и рядов Чебышева / С. Пашковский; пер. с пол. С. Н. Киро; под ред. В. И. Лебедева. – М. : Наука, 1983. – 384 с.
16. Gao, Q., Nie, C. B. An accurate and efficient Chebyshev expansion method for large-scale transient heat conduction problems // *Computers and Structures*. – 2021. – Vol. 249. – Article 106513. – DOI: 10.1016/j.compstruc.2021.106513.
17. Хартман, Ф. Обыкновенные дифференциальные уравнения / Ф. Хартман; пер. с англ. под ред. В. М. Алексеева. – М. : Мир, 1970. – 720 с.
18. Самарский А. А. Введение в численные методы. – М.: Наука, 1982. – 256 с.

© М. А. Новиков, В. В. Лисица, 2026