Multistage pseudo-spectral method (method of collocations) for the approximate solution of an ordinary differential equation of the first order

Konstantin P. Lovetskiy\textsuperscript{1},
Dmitry S. Kulyabov\textsuperscript{1,2}, Ali Weddeye Hissein\textsuperscript{1}

\textsuperscript{1}Peoples’ Friendship University of Russia (RUDN University),
6, Miklukho-Maklaya St., Moscow, 117198, Russian Federation
\textsuperscript{2}Joint Institute for Nuclear Research,
6, Joliot-Curie St., Dubna, Moscow Region, 141980, Russian Federation

(accepted: April 19, 2022)

Abstract. The classical pseudospectral collocation method based on the expansion of the solution in a basis of Chebyshev polynomials is considered. A new approach to constructing systems of linear algebraic equations for solving ordinary differential equations with variable coefficients and with initial (and/or boundary) conditions makes possible a significant simplification of the structure of matrices, reducing it to a diagonal form. The solution of the system is reduced to multiplying the matrix of values of the Chebyshev polynomials on the selected collocation grid by the vector of values of the function describing the given derivative at the collocation points. The subsequent multiplication of the obtained vector by the two-diagonal spectral matrix, ‘inverse’ with respect to the Chebyshev differentiation matrix, yields all the expansion coefficients of the sought solution except for the first one. This first coefficient is determined at the second stage based on a given initial (and/or boundary) condition. The novelty of the approach is to first select a class (set) of functions that satisfy the differential equation, using a stable and computationally simple method of interpolation (collocation) of the derivative of the future solution. Then the coefficients (except for the first one) of the expansion of the future solution are determined in terms of the calculated expansion coefficients of the derivative using the integration matrix. Finally, from this set of solutions only those that correspond to the given initial conditions are selected.

Key words and phrases: initial value problems, pseudo spectral collocation method, Chebyshev polynomials, Gauss–Lobatto sets, numerical stability

1. Introduction

Spectral methods are a class of methods used in applied mathematics and scientific computing to solve many differential equations numerically [1]–[4]. The main idea of the method is to represent the desired solution of a differential equation as a sum of certain ‘basis functions’ [5] (e.g., as an expansion into a sum in power functions — a Taylor series, or a sum of...
sinusoids, which is a Fourier series), and then calculate the coefficients in the sum to satisfy the differential equation in the best possible way.

Spectral and finite element methods are closely related and are based on the same ideas. The main difference between them is that spectral methods use nonzero basis functions over the entire domain, while finite element methods use nonzero basis functions only on small subdomains. In other words, spectral methods use a global approach, while finite element methods use a local approach. It is for this reason that spectral methods provide excellent convergence, their ‘exponential convergence’ being the fastest possible when the solution is smooth.

Spectral methods for the numerical solution of ordinary differential equations with given initial conditions are often reduced to solving a system of linear algebraic equations (SLAE), which includes both the initial conditions and conditions that ensure the fulfillment of differential relations [6]. However, a priori embedding of the initial (boundary) conditions into the system of linear equations leads to a significant increase in the filling of the matrices and, consequently, to the complication of the algorithm and method for solving the problem [7].

A more interesting approach is to select a basis that automatically takes into account the boundary conditions [1], [5], [6]. This is a frequently used trick when formulating the SLAE of the initial problem, and it reduces to taking into account the required initial/boundary conditions when creating the basis (a set of good basis functions-orthogonal, etc.) in a natural way, i.e., a basis is selected in which each basis function satisfies the initial conditions. The solution obtained using this approach is automatically sought in the class of functions satisfying the initial conditions. However, in this case it becomes much more difficult to work with new basis functions.

The novelty of the approach proposed by the authors is that first, a class (set) of functions that satisfy the differential equation is selected using a stable and computationally simple method of interpolation (collocation) of the derivative of the future solution. Then the coefficients (except for some) of the expansion of the future solution are determined in terms of the calculated expansion coefficients of the derivative using the integration matrix. Only after that, from this set of solutions those that correspond to the given initial conditions are selected.

Here we propose to divide the main problem into independent subproblems and to calculate the solution components in parts — separately those that determine the behavior of the derivative of the solution, and separately those that are determined by the boundary conditions. Thus, the problem is divided into two independent subproblems, each allowing stable and simple solution. The solution of the first problem in the simplest case is reduced to multiplying the vector of the right-hand side by the matrix of the Chebyshev functions values on the Gauss–Lobatto grid. At the next step, we solve the SLAE with a diagonal positive definite matrix and, multiplying the resulting vector on the left by the two-diagonal matrix, inverse (anti–derivative) with respect to the spectral Chebyshev matrix of differentiation, we obtain all the expansion coefficients of the desired solution, except for the first one. At the second, ‘most difficult’ stage, we determine the first coefficient of the expansion of the solution in terms of basis polynomials, solving a linear algebraic equation of the first order with respect to this coefficient.
2. Numerical solution of ordinary differential equations

Exact solution of a trivial ordinary differential equation for a given initial (boundary) condition

\[ y' = f(x), \quad x \geq x_0, \quad y(x) = y_0, \]

the right-hand side of which is independent of \( y \), can be presented in the form

\[ y_0 + \int_{x_0}^{x} f(\tau) d\tau. \]

Since the numerical methods for integrating functions are well developed from theoretical and practical points of view, it seems natural to apply them to the numerical solution of ordinary differential equations of general form

\[ y'(x) = f(x, y(x)), \quad x \geq x_0, \quad y(x_0) = y_0, \]

and this is exactly the fact that naturally explains the development and wide use of the methods of the Runge–Kutta type.

Usually, the method implies obtaining the solution in the interval \([x_0, x_0 + c_k h]\). The coefficients \( 0 \leq c_1 < c_2 < \ldots < c_n \leq 1 \) are chosen. Then, using the method of polynomial collocation, the solution is approximated by a polynomial \( p \) of the degree \( n \), which satisfies two types of conditions

- the initial condition: \( p(x_0) = y_0 \), and
- the differential equation, \( p'(x_k) = f(x_k, p(x_k)) \), at all the collocation points \( x_k = x_0 + c_k h \), \( k = 1, \ldots, n \).

Satisfying these \((n + 1)\) conditions allows calculating \((n + 1)\) coefficients of the expansion of the sought polynomial \( p \) of the degree \( n \).

Thus, the collocation methods are actually implicit Runge–Kutta methods [8].

It is important to note that to solve the approximation problem, it is not necessary to try solving Eq. (1) with simultaneous satisfaction of both the initial condition and the differential equation at the collocation points. In some cases, a fast and stable result can be achieved in two stages. First, to find those coefficients of the sought solution expansion that satisfy the differential equation at the collocation points. Then, to determine the deficient coefficients of the sought function expansion using the initial (final or intermediate) value.

3. Approximation of derivative. Cauchy problem

First, consider the problem of determining (recovering) a function from its derivative and some (one) additional condition. In this formulation, the problem naturally splits into two sub-problems:

- polynomial interpolation of the derivative (calculating the coefficients of the expansion of the derivative into a finite series in basis functions) and
- calculation of the coefficients of the required function by the initial (boundary, etc.) condition and the coefficients of the derivative expansion.

Without loss of generality, we assume that the domain of definition of the solution is the interval \([-1, 1]\).
Most often, the approximation of continuous functions is obtained by discarding the terms of the Chebyshev series, the magnitude of which is small \[9\], \[10\]. In contrast to the approximations obtained using other power series, the approximation in Chebyshev polynomials (having the property of being almost optimal) minimizes the number of terms required to approximate a function by polynomials with a given accuracy. This is also related to the property that the approximation based on the Chebyshev series turns out to be close to the best uniform approximation (among polynomials of the same degree), but easier to find. In addition, it allows avoiding the Gibbs effect with a reasonable choice of interpolation points.

Let us consider in more detail the problem of finding the derivative of the desired function, or rather the approximating polynomial \( p(x) \), satisfying condition (1) at a given number of points in the interval [\(-1, 1\)]. The pseudospectral (collocation) method [11] for solving the problem consists in representing the desired approximating function in the form of an expansion in a finite series in Chebyshev polynomials

\[
p(x) = \sum_{k=0}^{n} c_k T_k(x), \quad x \in [-1, 1]
\]

using the basis of Chebyshev polynomials of the first kind \( \{ T_k(x) \}_{k=0}^{\infty} \), defined in the Hilbert space of functions on the segment [\(-1, 1\)].

Let us differentiate the function (3). The derivative is expressed as

\[
p'(x) = \frac{d}{dx} \left( \sum_{k=0}^{n} c_k T_k(x) \right) = \sum_{k=0}^{n} c_k T'_k(x) = \sum_{k=0}^{n} b_k T_k(x), \quad x \in [-1, 1].
\]

Using the recurrence relations, which are satisfied by the Chebyshev polynomials of the first kind and their derivatives [3], [12] and equating the coefficients at the same polynomials in (4), we come [3] to the following dependence of the coefficients \( c_k \) on \( b_k \):

\[
\begin{bmatrix}
1 & 0 & -1/2 & 0 & \ldots & 0 & 0 \\
0 & 1/4 & 0 & -1/4 & \ldots & 0 & 0 \\
0 & 0 & 1/6 & 0 & \ldots & 0 & 0 \\
0 & 0 & 0 & 1/8 & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \ldots & 1 & 0 \\
0 & 0 & 0 & 0 & \ldots & 0 & 1/(2n)
\end{bmatrix}
\times
\begin{bmatrix}
b_0 \\
b_1 \\
b_2 \\
b_3 \\
\vdots \\
b_{n-2} \\
b_{n-1}
\end{bmatrix}
=
\begin{bmatrix}
c_1 \\
c_2 \\
c_3 \\
c_4 \\
\vdots \\
c_{n-1} \\
c_n
\end{bmatrix}.
\]

That is, the vector calculation of the coefficients \( \{ c_1, c_2, \ldots, c_n \} \) is the result of multiplying a simple tridiagonal matrix by a vector and it can be implemented by the following explicit formulas.
Thus, known the expansion coefficients \( b_k \) of the function \( f(x) \) of problem (1) in Chebyshev polynomials of the first kind, we can recover the last \( N \) expansion coefficients of the sought function in the same basis functions by formulas (2.1.3) from [3].

Therefore, the first part of the problem is to calculate the coefficients \( \{b_0, b_1, \ldots, b_n\} \) of the representation of the right-hand side of (1) on the interval \([-1, 1]\)

\[
\sum_{k=0}^{n-1} b_k T_k(x) = f(x).
\]

The collocation method consists in the selection of such coefficients \( \{b_0, b_1, \ldots, b_n\} \) of the expansion of the interpolation polynomial \( p'(x) \) that the following equalities are satisfied for the desired coefficients \( b_k, k = 0, 1, \ldots, n-1 \).

\[
\sum_{k=0}^{n-1} b_k T_k(x_j) = f(x_j), \quad j = 0, \ldots, n-1
\]

at the collocation points \( \{x_0, x_1, \ldots, x_n\} \).

The last statement is equivalent to the fact that the coefficients \( b_k, k = 0, \ldots, n \) must be a solution to the system of linear algebraic equations (7) of the collocation method. In matrix form, this can be written as

\[
Tb = f.
\]

The choice of collocation points should ensure the nondegeneracy of the system of Eqs. (7); for this it is sufficient that all grid points are different, and otherwise their choice is arbitrary, that is, the solution of system (7) on an arbitrary grid of the interval \([-1, 1]\) determines the required approximation. For an arbitrary grid, the matrix \( T \) is completely filled and the solution of such a system is rather laborious. To simplify the form of the matrix and speed up the search for the vector \( b \), we use the discrete orthogonality property of the Chebyshev matrix \( T \) on the Gauss–Lobatto grid. Consider the set \( x_j = \cos(j/n), j = 0, \ldots, n \) as collocation points. To further improve the properties of the system of linear equations, the solution of which will be the vector \( \{b_0, b_1, \ldots, b_n\} \), we multiply the first and last equations (7) by the factor \( 1/\sqrt{2} \). We obtain an equivalent ‘modified’ system with a new matrix \( \tilde{T} \) (instead of \( T \)) and a vector \( \tilde{f} \) instead of \( f \). The good thing about the new system is that it has the property of discrete ‘orthogonality’ and multiplying it on the left by the transposed \( \tilde{T}^T \) gives a diagonal matrix.
We transform system (8), multiplying it on the left by the transposed matrix $\tilde{T}^T$. As a result, we obtain a simple matrix equation with a diagonal matrix to determine the required expansion coefficients $\{b_0, b_1, \ldots, b_n\}$:

$$
\begin{bmatrix}
n & 0 & 0 & \ldots & 0 \\
0 & n/2 & 0 & \ldots & 0 \\
0 & 0 & n/2 & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \ldots & n
\end{bmatrix}
\begin{bmatrix}
b_0 \\
b_1 \\
b_2 \\
\vdots \\
b_n
\end{bmatrix} =
\begin{bmatrix}
\tilde{f}_0 \\
\tilde{f}_1 \\
\tilde{f}_2 \\
\vdots \\
\tilde{f}_n
\end{bmatrix}.
$$

(9)

Denoting by $(\tilde{f}_0, \tilde{f}_1, \ldots, \tilde{f}_{n-1}, \tilde{f}_n)^T$ the product of matrix $\tilde{T}^T$ by vector $(f_0/\sqrt{2}, f_1, \ldots, f_{n-1}, f_n/\sqrt{2})^T$ in the right-hand side of equation (9), we write down the required expansion coefficients of the derivative of the solution — the function $f(x)$ — in the explicit form

$$
\begin{aligned}
b_0 &= \frac{\tilde{f}_0}{n}, \\
b_1 &= \frac{2\tilde{f}_1}{n}, \\
b_2 &= \frac{2\tilde{f}_2}{n}, \\
\vdots & \\
b_n &= \frac{\tilde{f}_n}{n}.
\end{aligned}
$$

(10)

Consequently, relations (10), (6) uniquely determine the last $n$ coefficients $\{c_1, c_2, \ldots, c_n\}$ of the expansion of the sought function $p(x)$, and to determine one more coefficient $c_0$ it is necessary to involve at least one more additional condition. This can be both a boundary condition at the left or right end of the interval of consideration of a function, or a condition for the passage of the desired function through any given point within the interval of specifying the function.

That is, the considered method makes it possible to solve, depending on the type of the additional condition, both the Cauchy problem with initial conditions and problems with boundary conditions of a general form, requiring, for example, the use of the iterative shooting method [4].
In the case when the boundary condition is specified at the left end of the integration interval, the zero coefficient is determined from the equation

\[ c_0 + \sum_{k=1}^{n} c_k T_k(-1) = y_0 \]  

by the formula (taking into account that \( T_k(-1) = (-1)^k \) for any Chebyshev polynomial

\[ c_0 = y_0 - \sum_{k=1}^{n} c_k T_k(-1) = y_0 - \sum_{k=1}^{n} c_k (-1)^k. \]  

If the additional ‘boundary’ condition is specified at an arbitrary point of the integration interval, \( y_b = y(x_b) \), \( x_b \in [-1, 1] \), then the coefficient \( c_0 \) is determined by the formula

\[ c_0 = y_b - \sum_{k=1}^{n} c_k T_k(x_b). \]  

At the right-hand end of the integration interval \( y_r = y(1) \), \( x_r = 1 \), the Chebyshev polynomials of any order take the value equal to 1 (\( T_k(1) = 1 \)). Therefore, the coefficient \( c_0 \) is determined by the formula

\[ c_0 = y_r - \sum_{k=1}^{n} c_k T_k(x_r) = y_r - \sum_{k=1}^{n} c_k. \]  

4. Examples with simplest differential equations

Reconstructing a function from its derivative and a boundary condition. Comparison with the Runge–Kutta–Fehlberg method [13]

\[ \frac{dy}{dx} = f(x), \quad y(0) = y_0, \quad x \in [a, b]. \]

Let us compare the solutions obtained by the Runge–Kutta method (subroutine RKF45) and the solutions obtained as previously described.

Let us specify a grid in the interval \([a, b]\), calculated by the formula

\[ x_j = \frac{b - a}{2} \cos \left( \frac{j}{N - 1} \right) + \frac{b + a}{2}, \quad j = 0, 1, \ldots, N - 1, \]

and related to the chosen Gauss–Lobatto grid in the interval \([-1, 1]\). The number of grid points equals \( N \), i.e., to recover the function from the given derivative and additional condition by our method, only \( N \) calculations of the function (the right-hand side) are needed, and the recalculation of these values into the expansion coefficients in Chebyshev polynomials will require only \( 2N \) divisions and \( 2N \) additions.

To solve the Cauchy problem by the Runge–Kutta–Fehlberg method, we applied the RKF45 algorithm on each subinterval of the grid calculated above on \([a, b]\).
Algorithms are compared when looking for a solution to the simplest problem

\[ \frac{dy}{dx} = \cos(x), \quad y(0) = 0, \quad x \in [-\pi, \pi]. \]

The calculation carried out by the Runge–Kutta method with automatic control of accuracy (not worse than \(10^{-9}\)) required about 800 calculations of the function values over the entire interval.

For the two–stage method of separation of unknowns, the results of the deviation of the calculated values from the exact ones at the grid points are given in the table 1.

<table>
<thead>
<tr>
<th>Number of grid points</th>
<th>11</th>
<th>13</th>
<th>15</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum deviation</td>
<td>(&lt; 4 \cdot 10^{-7})</td>
<td>(&lt; 5 \cdot 10^{-9})</td>
<td>(&lt; 2 \cdot 10^{-13})</td>
<td>(&lt; 10^{-19})</td>
</tr>
</tbody>
</table>

Consider a few more model examples of solving the Cauchy problem, i.e., recovering functions from given derivatives and an initial condition. Functions from [14], in which the accuracy of calculating derivatives with the help of Chebyshev matrices of differentiation in physical space, were studied as model ones. The selected examples systematically illustrate the accuracy of calculating derivatives as a function of the number of approximation points (see the figure 1).

Four functions characterized by different smoothness are considered: \(|x^3|\), \(\exp(-x^{-2})\), \(1/(1 + x^2)\), and \(x^{10}\). The first function has the third derivative of bounded variation, the second function is smooth, but not analytical, the third one is analytical in the vicinity of \([-1, 1]\), and the fourth function is a polynomial. The accuracy of solutions obtained by us is by 1.5–3 orders of magnitude better than Trefethen’s solutions [14] when using a similar number of collocation points.

5. Discussion and conclusion

There are methods based mainly on the local approximation of the solution, which primarily use the initial approximation (boundary conditions) when solving differential equations. These are methods like Euler, Runge–Kutta method, etc. Other methods based on the approximation of the solution using global functions [global collocation methods — Mason, Boyd, Fornberg, Iserles, Townsend] are based on the construction of such systems of equations that simultaneously include both initial (boundary) conditions and conditions specifying the behavior of the derivatives of the desired solution.

In our study (within the framework of the pseudospectral collocation method), the problem is divided into two independent subproblems. The first is to select a set of solutions that satisfies the differential equation. However, it does not necessarily satisfy the initial (boundary) conditions. The choice of suitable bases for representing the solution in the form of an expansion
in polynomials, e.g., Jacobi ones, and grids taking into account the discrete orthogonality of the considered bases, makes it possible to use highly efficient algorithms.

Perhaps, the most promising from the point of view of the application of numerical methods is the use of a particular case of Jacobi polynomials — Chebyshev polynomials, as specific basis functions [15]. The Chebyshev polynomials provide an almost optimal approximation of the ODE solution, the ease of calculating the Gauss–Lobatto grid for using the discrete orthogonality condition, and three-term relations determining the ease of constructing differentiation and integration matrices of the sought solutions [16].

The initial (boundary) conditions are considered at the second stage of solving the original problem, which is actually reduced to solving a linear equation with one unknown coefficient.

The solution of the first problem is reduced to multiplying the matrix of values of the Chebyshev functions on the Gauss–Lobatto grid by the vector of values of the function that defines the right-hand side of the original differential equation to determine the expansion coefficients of the solution derivative. Further, the multiplication of the two-diagonal integration matrix [3] by the vector of coefficients yields all the coefficients of the desired solution, except for the first one. At the second stage, the use of the initial (boundary) condition makes it possible to determine the first coefficient of the solution expansion.
The approach based on dividing the problem of solving first-order ODEs into two subproblems seems to be very promising. The authors will continue to develop approaches and algorithms for the application of the multistage pseudospectral method for solving initial and boundary value problems with differential equations of higher orders.

Acknowledgments

This paper has been supported by the RUDN University Strategic Academic Leadership Program.

References


For citation:

Information about the authors:

**Lovetskiy, Konstantin P.** — Candidate of Physical and Mathematical Sciences, Assistant Professor of Department of Applied Probability and Informatics of Peoples’ Friendship University of Russia (RUDN University) (e-mail: lovetskiy-kp@rudn.ru, phone: +7(495)9522572, ORCID: https://orcid.org/0000-0002-3645-1060, ResearcherID: A-5725-2017, Scopus Author ID: 18634692900)

**Kulyabov, Dmitry S.** — Docent, Doctor of Sciences in Physics and Mathematics, Professor at the Department of Applied Probability and Informatics of Peoples’ Friendship University of Russia (RUDN University) (e-mail: kulyabov-ds@rudn.ru, phone: +7(495)9520250, ORCID: https://orcid.org/0000-0002-0877-7063, ResearcherID: I-3183-2013, Scopus Author ID: 35194130800)

**Hissein, Ali Weddeye** — student of Department of Applied Probability and Informatics of Peoples’ Friendship University of Russia (RUDN University) (e-mail: 1032209306@rudn.ru, phone: +7(977)4294833, ORCID: https://orcid.org/0000-0003-1100-4966, ResearcherID: AAD-7566-2022)
Многостадийный псевдоспектральный метод (метод коллокаций) приближенного решения обыкновенного дифференциального уравнения первого порядка

К. П. Ловецкий\textsuperscript{1}, Д. С. Кулябов\textsuperscript{1,2}, Али Уэддей Хиссен\textsuperscript{1}

\textsuperscript{1} Российский университет дружбы народов, ул. Миклухо-Маклая, д. 6, Москва, 117198, Россия
\textsuperscript{2} Объединённый институт ядерных исследований, ул. Жолио-Кюри, д. 6, Дубна, Московская обл., 141980, Россия

Аннотация. Рассмотрен классический псевдоспектральный метод коллокации, основанный на разложении решения по базису из полиномов Чебышева. Новый подход к формированию систем линейных алгебраических уравнений для решения обыкновенных дифференциальных уравнений с переменными коэффициентами и с начальными (и/или граничными) условиями позволяет значительно упростить структуру матриц, приводя её к диагональной форме. Решение системы сводится к умножению матрицы значений полиномов Чебышева на выбранной сетке коллокации на вектор значений функции, описывающей заданную производную в точках коллокации. Следующее за этой операцией умножение полученного вектора на двухдиагональную спектральную «обратную» по отношению к матрице дифференцирования Чебышева приводит к получению всех коэффициентов разложения искомого решения за исключением первого. Этот первый коэффициент определяется на втором этапе исходя из заданного начального (и/или граничного) условия. Новизна подхода заключается в том, чтобы сначала выделить класс (множество) функций, удовлетворяющих дифференциальному уравнению, с помощью устойчивого и простого с вычислительной точки зрения метода интерполяции (коллокации) производной будущего решения. Затем рассчитать коэффициенты (кроме первого) разложения будущего решения по вычисленным коэффициентам разложения производной с помощью матрицы интегрирования. И лишь после этого выделить из этого множества решений те, которые соответствуют заданным начальным условиям.

Ключевые слова: начальные задачи, метод псевдоспектральных коллокаций, многочлены Чебышева, множества Гаусса–Лобатто, численная устойчивость