

SOLUTION OF THE CAUCHY PROBLEM FOR ORDINARY DIFFERENTIAL EQUATIONS USING THE COLLOCATION AND LEAST SQUARES METHOD WITH THE PADE APPROXIMATION

V.P. Shapeev^{1,2}

¹Khristianovich Institute of Theoretical and Applied Mechanics SB RAS,
Novosibirsk, Russian Federation

²Novosibirsk State University, Novosibirsk, Russian Federation

A new method for solving the Cauchy problem for an ordinary differential equation is proposed and implemented using the collocation and least squares method of increased accuracy. It is based on the derivation of an approximate nonlinear equation by a multipoint approximation of the problem under consideration. An approximate solution of the problem in the form of the Pade approximation is reduced to an iterative solution of the linear least squares problem with respect to the coefficients of the desired rational function. In the case of nonlinear differential equations, their preliminary linearization is applied. A significant superiority in accuracy of the method proposed in the paper for solving the problem over the accuracy of the NDSolve procedure in the Mathematica system is shown. The solution of a specific example shows the superiority in accuracy of the proposed method over the fourth-order Runge–Kutta method. Examples of solving the Cauchy problem for linear and non-linear equations with an accuracy close to the value of rounding errors during operations on a computer with numbers in the double format are given. It is shown that the accuracy of solving the problem essentially depends on the complexity of the behavior of the values of the right-hand side of the equation on a given interval. An example of constructing a spline from pieces of Pade approximants on partial segments into which a given segment is divided is given in the case when it is necessary to improve the accuracy of the solution.

Keywords: Cauchy problem; ordinary differential equation; Pade approximation; collocation and least squares method; high accuracy; Mathematica System.

Introduction

The fractional rational functions discovered by Henri Pade aroused the interest of mathematicians in studying their properties primarily from the point of view of the theory of functions. A large number of works are devoted to this topic [1 – 7]. One of the most important properties of fractional rational functions is their applicability to find high-precision approximations of Pade approximation functions, which can be much more accurate than polynomial ones. For them, we will use the generally accepted designation $[L/M]$ and PA together. In the notation $[L/M]$, the value of L is the degree of the polynomial in the numerator, and M is the degree of the polynomial in the denominator of fractional rational functions.

For objective reasons, the use of Pade approximations for solving computational mathematics problems is inferior in the number of cases of using other types of approximation. However, recently they were used relatively more often for approximate solutions of various problems of numerical analysis. In this analysis, one of the directions is the application of the Pade approximation to the solution to ordinary differential equations (ODEs) [8 – 10]. One of the reasons for the relatively rare use of $[L/M]$ in this direction

is the difficulty of finding it. Both in the case of the use of fractional rational functions in solving problems of approximation of given functions, and in some works devoted to solving the Cauchy problem for ODEs, the connection of the PA with chain fractions is used. In this approach to solving differential equations, the derivation of complex formulas is required to find a numerical solution to the problem. This fact makes it difficult to apply it in applications, which are indirectly indicated in [9] when discussing the content of the book [8]. Another method, as in the case of constructing PA analytical functions, is based on using their connection with series expansions [10]. The accuracy of the solution obtained by applying this property is limited by the radius of convergence of the series. If there is a need for a global solution to the problem, there may be a need to continue the solution from a circle with this radius. Therefore, in [10], if such a continuation is necessary, an additional requirement is imposed on the condition of the problem: “if there is a need for the practical implementation of the proposed algorithm, (the continuation) must be algorithmically specified”. The practical application of such an approach is no less cumbersome than the method using chain fractions and unproductive for obtaining quantitative results of solving the Cauchy problem.

When finding the approximation of functions using the PA, an approach based on multipoint approximations [4] is used in combination with the least squares method, which has not yet found application in solving differential equations. In the least squares and collocation method collocation and least squares (CLS) method, the approximate solution of differential [11 – 16] and integral [17] equations is reduced to solving redefined systems of linear algebraic equations (SLAE). At the same time, the use of the least squares method in it allows minimizing the error functional of the corresponding SLAE and significantly expands the possibilities of using the collocation method with polynomial approximation to solve these equations. Currently, the CLS method has solved with improved accuracy a number of problems for elliptic, parabolic and hyperbolic equations that are used to model various physical processes.

The purpose of this work is to show the possibility of a high-precision solution of the Cauchy problem for a first-order ODE by the iterative solution of the equation obtained here by its multipoint PA, to show the quantitative characteristics of the solution process and the found $[L/M]$, as well as to demonstrate the results of combining this approach with the least squares method for solving a SLAE with a rectangular matrix, which is obtained in the result of a multipoint approximation of the Cauchy problem for the ODE. It seems that in the available small number of works on solving the problem under consideration, these issues are not sufficiently disclosed to attract the attention of computational mathematics specialists to it. The approach to solving the problem considered here is quite easily implemented in a computer program, which facilitates its use for solving specific problems, obtaining their quantitative characteristics and evaluating the accuracy of numerical results.

Here, in numerical experiments on solving the Cauchy problem for the ODE, it was observed that the conditionality of the SLAE obtained as a result of the approximation of the problem depends not only on its size, but also on the ratio of the number of equations in it to the number of desired coefficients $[L/M]$. As is known, the possibility and accuracy of its solution by one method or another depends on the conditionality of the SLAE. So it depends on the conditionality of the SLAE whether an improved accuracy of the approximate solution of the problem in the form of $[L/M]$ can be achieved.

This position takes place in various problems of computational mathematics both in the search for polynomial approximants and in the search for Pade approximations by various methods. The need for high accuracy of intermediate calculations when finding $[L/M]$ to approximate one-dimensional functions using their connection with Taylor series expansion is already mentioned in the introduction of the book [1]. This may indicate a poor conditionality of the problem of finding $[L/M]$. However, questions about the stability and the accumulation of rounding errors of existing methods for finding Pade approximations due to their algorithmic complexity have not yet been discussed in the scientific literature.

1. Problem Statement and Method of Its Solution

Let the Cauchy problem is to be solved:

$$dy(x)/dx = f(x, y), \quad x \in [a, b]; \quad y(\xi) = y_0, \quad \xi \in [a, b]. \quad (1)$$

Let the function $f(x, y)$ be discretely given at n grid nodes on the segment $[a, b]$, and the Cauchy condition y_0 is given at some point ξ on $[a, b]$. In the proposed algorithm for solving the problem, any point of the segment $[a, b]$ can be formally and often practically set as ξ , provided that there is a solution to such a problem.

Let us search for an approximate solution of the problem (1) $y(x)$ in the form $[L/M]$:

$$y(x) = \frac{a_0 + a_1x + \dots + a_Lx^L}{b_0 + b_1x + \dots + b_Mx^M} = \frac{N(x)}{D(x)}, \quad (2)$$

in which, for its unambiguous definition, we put $b_0 = 1$. On $[a, b]$ we define a grid with n nodes satisfying the inequality $L + M + 1 < n$ by choosing the number of grid nodes. We write SLAE to find $L + M + 1$ indeterminate coefficients $a_0, \dots, a_L, b_1, \dots, b_M$ from the requirement of approximation by the function $[L/M]$ of the Cauchy equation and condition (1) by collocation in the nodes of the grid.

After substituting (2) into the differential equation (1), we obtain

$$\frac{dy(x)}{dx} = \left(D(x) \frac{dN(x)}{dx} - N(x) \frac{dD(x)}{dx} \right) \frac{1}{D^2(x)}. \quad (3)$$

Taking into account (2), the expression (1) can be written as:

$$D(x) \frac{dy(x)}{dx} = \frac{dN(x)}{dx} - \frac{N(x)}{D(x)} \frac{dD(x)}{dx} = \frac{dN(x)}{dx} - y(x) \frac{dD(x)}{dx}. \quad (4)$$

The expression (1) is nonlinear with respect to the desired coefficients $[L/M]$. In order to reduce the search for an approximate solution of the problem to a SLAE solution, (4) can be solved iteratively. Let us rewrite it in the form

$$D_k(x)f(x) = \frac{dN_k(x)}{dx} - g_{k-1}(x) \frac{dD_k(x)}{dx}, \quad k = 1, 2, 3, \dots \quad (5)$$

and we assume in the equation at each k -th iteration $g_{k-1}(x)$ the solution of the problem obtained at the $(k - 1)$ -th iteration. In the case of convergent iterations (5) in the limit goes to (4). We approximate the Cauchy condition at a given point ξ .

$$N_k(\xi) - y_0 D_k(\xi) = 0. \quad (6)$$

Since $f(x, y)$ and $g_{k-1}(x)$ are known after the $(k - 1)$ -th iteration, then (5) and (6) are linear with respect to the coefficients of the polynomials $N_k(x)$ and $D_k(x)$. As a result of collocation of relations (5), (6) in the grid nodes, a SLAE with a matrix with numerical elements is obtained. In the iterative process of its solution, at each iteration, in the case of their convergence, it determines a refined approximation to the solution of the problem. Let us denote it as $Ax = b$, where A is a rectangular matrix. The number of necessary iterations depends on the required accuracy of the solution chosen initial approximation of $g_0(x)$ and, as shown below in solving the problems given here, on the type of equation (1), on which an important value depends in the numerical solution of the problem – the condition number of the obtained SLAE. In the numerical experiments carried out to solve the Cauchy problem with different complexity of equations using the proposed algorithm, it was found that in many cases it converges quite quickly, even if the initial approximation was assumed to be a constant over the entire segment $[a, b]$. The number of iterations when searching for $[L/M]$ can be reduced if we take $[L_1/M_1]$ as $g_0(x)$, found using the method proposed here with the values $L_1 < L$ and $M_1 < M$. In the case when the right side of the equation in problem (1) is an analytical function of a simple form, without large gradients or any significant singularities on $[a, b]$, using calculations on a computer with numbers in double format allows using the proposed algorithm to obtain a solution to problem (1) with an error value in the interval $(10^{-15}, 10^{-14})$. At the same time, a high-precision Pade approximation can be found with a small number of coefficients $[L/M]$ – the degree of its freedom.

The algorithm for solving the Cauchy problem for the ODE using the relations (5) and (6) is implemented here in the form of a program in the Mathematica system [18], which for brevity we will call the APODE procedure. Mathematica was originally created to perform various transformations of expressions and formulas on a computer, written in a symbolic form close to the form used by mathematicians. Subsequently, a large library of procedures for the numerical solution of various mathematical problems appeared in it. The procedure of the Mathematica NDSolve system solves numerically on a given interval the symbolically written Cauchy problem for the ODE. The application of the evaluate procedure in Mathematica to the numerical solution obtained by her results in some approximation of the solution, which allows calculating the values of the solution at an arbitrary point of this segment without repeated calls to NDSolve. These two procedures of the Mathematica system are used here as a convenient tool at the stage of verification of the APODE procedure and variants of the program implementing it. It should be noted at once that the solutions obtained using NDSolve for a significant number of Cauchy problems in the performance of this work, some of which are given here, have errors in the interval $(10^{-6}, 10^{-5})$. It does not exceed the accuracy of the representation of numbers on a computer in float (C++) format, but it can be sufficient for solving many applied problems. Note that the user of the NDSolve procedure in the Mathematica system, given a specific Cauchy problem, has no way to influence the accuracy of its solution. It lacks parameters that determine the accuracy of the solution.

The input data of the APODE procedure, which the user selects and sets, are the values of the parameters L , M and n , the number of collocation points (the number of grid nodes on the segment $[a, b]$). For the convenience of the user, there are two variants of the condition for terminating iterations: according to the specified number of iterations or upon reaching a given value of the norm of the difference of approximate solutions

obtained at the current k -th and the previous $(k - 1)$ -th iterations. In the second iteration termination criterion, the sought $[L/M]$ is conditionally considered as a vector whose components are its coefficients, and the Euclidean norm of the vector is taken as the norm of the difference between the two approximations. For a visual demonstration of the capabilities of the APODE procedure, the following results of numerical experiments give the error values of the solutions of the ODE obtained by its application. At the same time, all calculations and comparisons of the numerical algorithm proposed in this paper for solving the Cauchy problem for ODE with other methods are carried out on a uniform grid, which is far from the best for obtaining high accuracy results, which can be more accurate when using adaptive grids.

2. Numerical Experiments

To verify the APODE procedure and the program implementing it, an example was first selected that had a known numerical solution with high accuracy. At the same time, the errors of the solutions obtained by the application APODE and NDSolve were compared.

Example 1. Let us consider the Cauchy problem

$$y'(t) = \sin(t)/t, \quad t \in [0, 7, 5], \quad y(0) = 1. \quad (7)$$

Its solution $Si(x)$ is an integral sine for which there is no (finite) exact formula expressed in terms of elementary functions. An appeal in the Mathematica system to the library function `SinIntegral[x]` gives its values with 16 decimal places. Note that in this example, when referring to the NDSolve procedure, the point $x = 0$ cannot be specified as the point for setting the Cauchy condition, since the right side of the equation in it has a singularity, and in programming systems the value $\sin(0)/0$ is not calculated. An attempt to calculate the values of $\sin(0)/0$ on the computer leads to an automatic termination of calculations by the computer – to an "Automatic computer shutdown". Calculation of limits is not provided in NDSolve. In order for the formulation of the problem solved by the two methods being compared to be the same in both cases, the segment $[\epsilon, 7, 5]$ was first taken, on which a grid with n nodes was set. However, it is essential that as a Cauchy condition in APODE, you can explicitly set $y(0) = 1$ and use the collocation method to include this condition in the SLAE defining the desired $[L/M]$. Table 1 shows part of the results at $\epsilon = 10^{-15}$, the value of which is close to the rounding errors in the arithmetic of numbers in double format and practically does not distort the error values of solutions. The solution of the problem by the APODE procedure under the condition $y(0) = 1$ differed from the solution in the case $y(\epsilon) = 1$ by an order of magnitude ϵ .

It should be noted that finding a solution to the Cauchy problem by constructing $[L/M]$ of improved accuracy reduces to solving a poorly conditioned SLAE even in cases of problems with the right side of the equations $f(x, y)$ of a relatively simple form with regular behavior of values on the segment under consideration. Here, the condition number of $cond(A)$ matrices was calculated in the spectral norm using Mathematica library procedures. A QR decomposition of its matrix was used using an orthogonal matrix Q when solving SLAE. The solution of the redefined SLAE in this way achieves the minimum functional of the residual of its equations, as well as the solution of the system $A^T Ax = A^T b$. However, the first method leads to a solution of a SLAE of the form

Table 1

The error $[L/M]$ of problem (7) solution
by the APODE procedure at various parameter values

	L	M	n	it	$err([L/M])$	$cond(R)$
№ 1	3	4	8	40	$9,876 \times 10^{-4}$	11019, 4
№ 2	3	4	16	5	$1,778 \times 10^{-4}$	9219, 9
№ 3	3	4	16	20	$1,331 \times 10^{-5}$	9193, 8
№ 4	9	10	20	20	$4,201 \times 10^{-13}$	$2,7 \times 10^{14}$
№ 5	9	10	20	40	$4,363 \times 10^{-13}$	$5,1 \times 10^{14}$
№ 6	24	25	81	40	$1,554 \times 10^{-15}$	$1,6 \times 10^{26}$

$Rx = Q^T b$ with a condition number $cond(R) = cond(A)$. It has a significant advantage over the second one when solving poorly conditioned SLAEs, since $cond(A^T A) = (cond(A))^2$ and can often be significantly larger than $cond(A)$. In Table 1, the number of iterations is written out in the column with the id it , and in the column $err([L/M])$ – the values of the error APODE. The error of the solutions and their convergence were estimated here and further in the uniform norm $\|\cdot\|_C$ calculated on a certain set of points G uniformly located on the segment of the solution of the problem. In this example, using the solutions obtained, the values $\max_{x \in G} |NDSolve - Si(x)|$ и $\max_{x \in G} |APODE - Si(x)|$ were calculated – the errors of the NDSolve and APODE procedures, respectively. The error of NDSolve is fixed and turned out to be equal to 7.154×10^{-7} . The analysis of Table 1 shows that even with small values of L, M, n, it , the accuracy of the APODE procedure significantly exceeds the accuracy of NDSolve and with increasing values of these parameters approaches the accuracy of the representation of real numbers in the double format. At the same time, $cond(R)$ grows strongly, and its calculation in the above way in Mathematica becomes unstable. When solving such SLAE, messages appear in the Mathematica system that they are poorly conditioned. When comparing the first and second rows of Table 1, it can be seen that the accuracy of the solution in the second row is better than in the first, although in the second case there were fewer iterations. This suggests that in order to obtain high accuracy of the solution, it is important to use multipoint approximation in combination with the least squares method. To do this, it is necessary that the number of collocation points exceeds the degree of arbitrariness of the desired $[L/M]$. This fact took place in all the calculations carried out in this work. In addition, the influence of the number of collocation points on the error $[L/M]$ is traced.

Similar results and comparisons on the accuracy of problem solutions by these procedures are obtained in the case of the above simple right sides of the equations. For example, the solution of the problem $y'(t) = 2y(t)$, $t \in [0, 1]$, $y(0) = 1$, obtained using APODE at $L = 14, M = 15, n = 46$, after eighteen iterations reaches an accuracy of the order of 10^{-15} , and the solution obtained by NDSolve is only $4,69 \times 10^{-6}$.

Example 2. Consider a test problem

$$y'(x) = 4x \cos(\pi/6 + 4x) + 4 \sin(\pi/4 - x) + \sin(\pi + 4x), \quad y(0) = 2\sqrt{2}, \quad x \in [0, 2\pi] \quad (8)$$

with the complex behavior of the right side of the equation on the segment under

consideration and with the corresponding behavior of the known exact solution

$$y(x) = 4 \cos(\pi/4 - x) + x \sin(\pi/6 + 4x). \quad (9)$$

The right side of the equation and its solution have about ten local extremes on the segment of the solution of the problem due to the difference in the periods of periodic functions in them. The gradient modulus of the right part reaches a value of the order of 10^2 . Fig. 1a shows a graph of values on the segment $[0, 2]$ of the right side of equation (8) and Fig. 1b shows a graph of the PA solution of the problem (8). This problem presents some difficulty when trying to obtain a solution of improved accuracy by numerical methods. Using the NDSolve procedure, a solution with an error of $3,73 \times 10^{-5}$ was obtained. Using APODE for 22 iterations, $[29/30]$ was obtained, which has an error of $4,83 \times 10^{-10}$. Note that with a fixed format for the representation of numbers in a computer program, in cases of examples with complex behavior of the right side of the equations, the possibility of improving the accuracy of solutions by increasing the values of the parameters L, M, n is limited. With their growth, condition number of SLAE increases and the accumulation of rounding error increases when solving it accordingly. One of the possibilities of improving the accuracy of solving the Cauchy problem by the APODE procedure is to construct a solution in the form of a spline from pieces of $[L/M]$ on partial segments into which a given segment is divided. At the same time, when constructing a continuous spline on each partial segment in the Cauchy condition, it is assumed to take the value of the solution obtained at the end of the previous partial segment. Here the initial segment $[0, 2]$ was divided into only 4 parts, and $[29/30]$ was found on each of them. In the first partial segment, the spline had an error of $3,29 \times 10^{-14}$, in subsequent segments it monotonically increased and in the last segment it turned out to be equal to $1,88 \times 10^{-11}$. The spline of the pieces $[29/30]$ is more accurate than the monofunctional approximant $[29/30]$ of the solution of the problem by more than 25 times. The question of the best way to split a segment in terms of the magnitude of the error of the spline was not considered here.

Fig. 2a shows a graph of the first link of the constructed spline (in this figure, to save space, the origin of coordinates along the ordinate axis is shifted compared to figure 1b, and in 2b – a graph of its error. It is characteristic that the error graph turned out to be a solid “veil” of high-frequency oscillations. This, firstly, corresponds to the fact that the process of rounding the values of arithmetic calculations on a computer has a random character, and, secondly, the difference between the amplitude of the oscillations and the accuracy of the representation of numbers in the double format indicates the value of the SLAE conditioning number, the solution of which is obtained $[29/30]$. As is known, in the magnitude of the error of the SLAE solution, the condition number stands as a multiplier before the magnitude of the perturbations of the input data of the problem: the values of the elements of the SLAE matrix and the components of the vector of its right side. Here, the reason for their perturbations are the errors of rounding numbers on the computer in the process of solving SLAE.

Example 3. The possibilities of the proposed approach and numerical algorithm are also tested on solving nonlinear equations. Here we present the results obtained by solving the Riccati equation arising in various applications

$$y'(x) = a(x)y^2(x) + b(x)y(x) + c(x), \quad (10)$$

which was solved with different types of its coefficients. The solution to equation (10) can only be written in particular cases in the form of a finite formula expressed in terms of

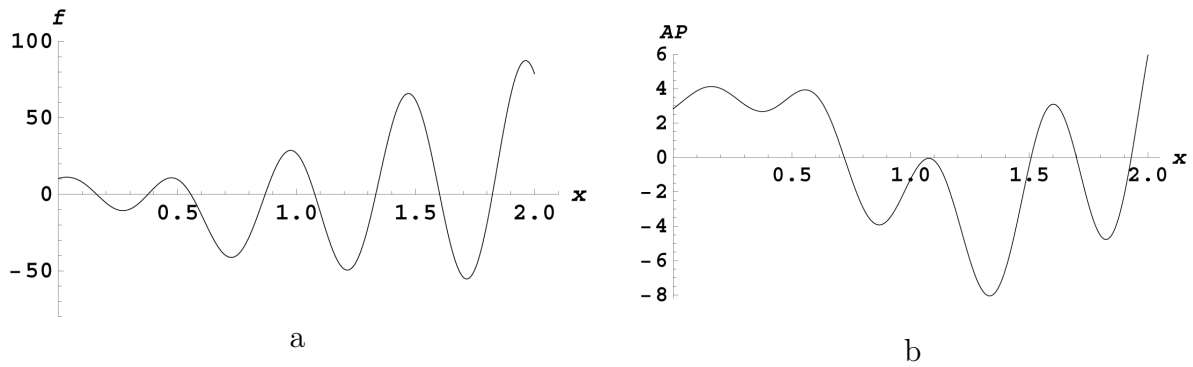


Fig. 1. Graph of the right side of equation (8) (a) and graph [20/30] are the solutions of the problem (8) on the segment $[0, 2]$ (b)

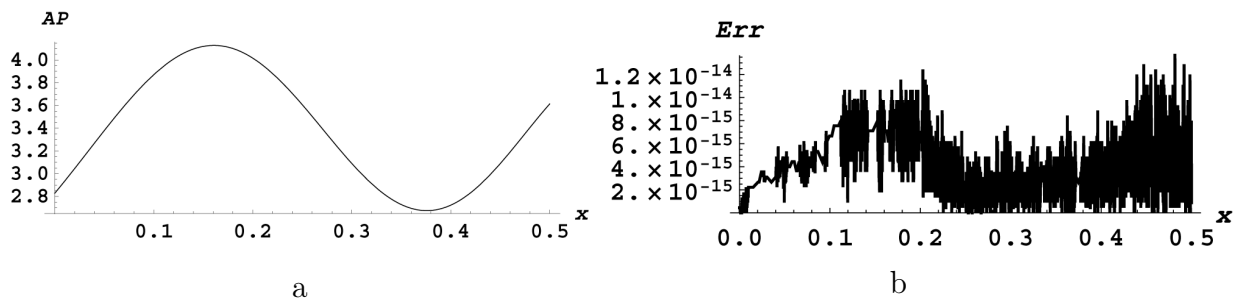


Fig. 2. Graph of the first link of the spline from pieces [20/30] (a) and the error graph of the first link of the spline [20/30] (b)

elementary functions. When finding an approximation in the case of the iterative solution of the Cauchy problem for (10), different ways of linearization of the summand $a(x)y^2(x)$ in the right side of the equation. In the first case, it was written in the form $a(x)y_k y_{k-1}$, in which the value of y_{k-1} in the approximate equation (5) was taken from the previous $(k - 1)$ iteration, and y_k was included in the SLAE with the coefficients sought on the k -th iteration. In the second case, a simple linearization method was used: the value of the term $a(x)y^2(x)$ was taken entirely from the $(k - 1)$ -th iteration. In the third case, when the summand was linearized by Newton, the iterations in the examples discussed here converged monotonically with a decrease in the discrepancy of the equation on the approximate solution with an increase in the iteration number and faster than in the other two cases. In the second case, the discrepancy decreased more slowly and the SLAE was worse than in the third case. In the first case, the rate of convergence of iterations and the value of the SLAE condition number were intermediate with respect to the other two cases.

Example 4. First, to verify the APODE, and an approximate assessment of its accuracy, consider the problem

$$y'(x) = y^2(x) - 2y(x) \exp(x) + \exp(2x) + \exp(x), \quad x \in [0, 1], \quad y(0) = 0 \quad (11)$$

with a known exact solution $y(x) = \exp(x) - 1/(x + 1)$. Equation (11) on the solution obtained using the NDSolve procedure has a discrepancy of 2.992×10^{-4} . And substituting into equation (11) the approximate solution of the problem obtained by the APODE procedure [11/12] with 36 collocation points and an initial approximation equal to one over the entire segment $[0, 1]$ gave a discrepancy of $5,063 \times 10^{-14}$. Knowledge of the exact

solution allows us to assess the difference in the possibility of achieving the improved accuracy of solving the problem by the APODE procedure and the Runge–Kutta high accuracy scheme of the 4-th order (RK-4). In numerical experiments for solving problem (11) on the segments $[0, 1]$ and $[0, 3]$, the error behavior was observed depending on the values of various parameters of the compared methods. With sequential grinding of the grid step, a decrease in the error of the RK-4 method was first observed until it reached the minimum value at some sufficiently small grid step. After that, it increased with a further decrease in the step. For example, when solving equation (11) on the segment $[0, 3]$ on a grid with 4751 nodes, after reaching the error value of $1,306 \times 10^{-12}$, with further grinding of the step, it increased as a result of the accumulation of the rounding error due to the large number of arithmetic operations performed. At the same time, for simplicity, the error of the RK-4 method was calculated only at the point $x = 3$, i.e. its maximum value was not searched for over the entire segment. As a result of solving the problem on the same segment, the APODE procedure obtained $[7/8]$ with the number of nodes $n = 22$, which has a maximum error of $1,145 \times 10^{-13}$ on the entire segment. A comparison of the results of solving the problem on the segment $[0, 1]$ showed that the APODE procedure gave a result a decimal order more accurate than RK-4. An important circumstance is that in an effort to achieve high accuracy of the solution by the RK-4 method, the calculator receives a large array of numerical values, the storage and use of which are inconvenient for him. At the same time, the time spent to obtain this array and the time spent building the corresponding $[L/M]$ on the computer differ by more than two decimal orders in favor of the second. As a result of consideration of some other examples of the Cauchy problem for ODEs with known exact solutions, similar results were obtained, indicating an advantage in accuracy of the APODE procedure over the RK-4 method. Note simultaneously that as a result of the application of the PA, the solution of the Cauchy problem for the ODE is obtained in the form of a finite analytical expression, which, with a small number of arithmetic operations, allows it to be calculated with high accuracy at any point of a given segment. Applying many other numerical methods to solve the Cauchy problem, we obtain a discrete solution at the nodes of a certain grid. If it is necessary to calculate it at the intermediate points of the segment, it is necessary to apply an additional approximate computational procedure, for example, interpolation, which introduces an additional error into it. This is another very important advantage of the PA in solving ODES over many other numerical methods.

Example 5. When solving the problem

$$y'(x) = y^2(x) - 6x^2, \quad x \in [0, 1], \quad y(0) = 1, \quad (12)$$

the initial approximation, as in most other problems solved here, was equal to 1 for the entire segment. Due to the lack of an exact solution, in this example, a difference in the values of the approximants of the solution $[11/12]$ in a uniform norm obtained on two consecutive iterations was observed. The number of collocation points was equal 40. As the number of iterations grew, the solution converged monotonically. In addition, the convergence of the residual norm of the equation was observed when substituting the obtained $[11/12]$ into it, which also monotonically tended to zero and at the 51-st iteration was equal to $1,332 \times 10^{-14}$. At the same time, $[11/12]$ obtained at the 51-st and 52-nd iterations completely matched all the coefficients, both in their numerators and denominators. In the convergent sequence of an iterative solution, there is a correlation

between its error and the discrepancy. This numerical experiment gives an approximate idea of the behavior of the error on the sequence of iterations of the approximate solution by the APODE procedure in cases where it is not possible to calculate it in the absence of a known exact solution. When solving the problem by the NDSolve procedure, the maximum of the residual of equation (12) was $2,99 \times 10^{-4}$. This indicates the relatively low accuracy of NDSolve when compared with APODE.

The examples of solving nonlinear equations given here show that Newton's preliminary linearization of nonlinear equations expands the classes of ODEs for which the application of the procedure proposed here allows iteratively finding high-precision $[L/M]$ solutions to the Cauchy problem.

Conclusion

Using the collocation and least squares method, an iterative procedure for constructing an $[L/M]$, the approximation of the solution of the Cauchy problem for the ODE is proposed and implemented in the Mathematica system. The construction of $[L/M]$ is reduced to solving a linear least squares problem for SLAE obtained by approximating the original differential problem. The possibility of using this procedure to iteratively find high-precision $[L/M]$ solutions of the Cauchy problem for linear and nonlinear ODEs is shown. At the same time, it has an accuracy superiority over the fourth-order Runge–Kutta method and a significant superiority over NDSolve, the standard procedure of the Mathematica system for representing the solution of the Cauchy problem as an approximant. The dependence of the accuracy of solutions on the complexity of the form of the right-hand side of the equations, on the degree of arbitrariness of the sought $[L/M]$ and on the of the SLAEs defining them is shown. In the case of a nonlinear equation, the possibility of constructing $[L/M]$ after preliminary linearization of the equation in different ways is tested. In the cases considered, Newton linearization proved to be the most effective. An example of constructing a spline from pieces of $[L/M]$ on partial intervals of a given segment of the solution of the Cauchy problem is given. It is shown that the spline of the pieces $[L/M]$ approximates the solution much more accurately in comparison with the monofunctional $[L/M]$ constructed over the entire segment.

Acknowledgments. *The research was supported by the Russian Science Foundation grant no. 23-21-00499.*

References

1. Baker G.A.Jr. *Essentials of Pade Approximants*. New York, Academic Press, 1975.
2. Gonchar A.A., Rakhmanov E.A. On the Convergence of Joint Pade Approximations for Systems of Markov Type Functions. *Proceedings of the Steklov Institute of Mathematics*, 1981, vol. 157, pp. 31–50.
3. Gonchar A.A., Rakhmanov E.A. Equilibrium Distributions and the Rate of Rational Approximation of an Analytical Function. *Mathematics of the USSR-Sbornik*, 1989, vol. 62, no. 2, pp. 305-348. DOI: 10.1070/SM1989v062n02ABEH003242
4. Gonchar A.A., Novikova N.N., Khenkin G.M. Multipoint of Pade Approximants. *Matematicheskyy Sbornik [Mathematical Collection]*, 1996, vol. 187, no. 12, pp. 57–86. (in Russian)

5. Suetin S.P. Pade Approximations and Effective Analytical Continuation of a Power Series. *Russian Mathematical Surveys*, 2002, vol. 57, no. 1, pp. 43–141. DOI: 10.1070/RM2002v057n01ABEH000475
6. Gonchar A.A. Rational Approximations of Analytic Functions. *Proceedings of the Steklov Institute of Mathematics*, 2011, vol. 257, no. 2, pp. 44–57. DOI: 10.1134/S0081543811030047
7. Aptekarev A.I., Buslaev V.I., Martines-Finkelstein A., Suetin S.P. Pade Approximations, Continuous Fractions and Orthogonal Polynomials. *Russian Mathematical Survey*, 2011, vol. 66, no. 6, pp. 1049–1131. DOI: 10.1070/RM2011v066n06ABEH004770
8. Khovansky A.N. *The Application of Continued Fractions and their Generalizations to Problems in Approximation Theory*. P. Noordhoff, Groivngen, the Netherlands, 1963.
9. Velichko I.G., Tkachenko I.G., Balabanova V.V. Application of the Method of Continued Fractions to Obtain Approximations of the Case Solutions of Cauchy Problems for Differential Equations of the First Order. *Bulletin of Science and Education in North-West Russia*, 2015, vol. 1, no. 3, pp. 1–9. (in Russian)
10. Vishnevsky V.E., Zubov A.V., Ivanova O.A. [Pade Approximation of the Solution of the Cauchy Problem]. *Vestnik Sankt-Peterburgskogo Universiteta*, 2012, no. 4, pp. 3–17. (in Russian)
11. Isaev V.I., Shapeev V.P. High-Accuracy Versions of the Collocations and Least Squares Method for the Numerical Solution of the Navier–Stokes Equations. *Computational Mathematics and Mathematical Physics*, 2010, vol. 50, no. 10, pp. 1670–1681. DOI: 10.1134/S0965542510100040
12. Shapeev V.P., Vorozhtsov E.V. CAS Application to the Construction of the Collocations and Least Residuals Method for the Solution of the Burgers and Korteweg-de Vries–Burgers Equations. *International Workshop on Computer Algebra in Scientific Computing*, Warsaw, 2014, vol. 8660, pp. 432–446. DOI: 10.1007/978-3-319-10515-4-31.
13. Shapeev V.P., Belyaev V.A. [Variants of the Collocation Method and the Least Residuals of Increased Accuracy in a Region with a Curvilinear Boundary]. *Vychislitel'nye Tekhnologii* [Computational Technologies], 2016, vol. 21, no 5. pp. 95–110. (in Russian)
14. Belyaev V.A., Shapeev V.P. Versions of the Collocation and Least Squares Method for Solving Biharmonic Equations in Non-Canonical Domains. *AIP Conference Proceedings*, 2017, vol. 1893, no. 1, article ID: 030102. DOI: 10.1063/1.5007560.
15. Shapeev V.P., Belyaev V.A., Golushko S.K., Idimeshev S.V. New Possibilities and Applications of the Least Squares Collocation Method. *EPJ Web of Conferences*, 2018, vol. 173, article ID: 01012. DOI: 10.1051/epjconf/201817301012
16. Vorozhtsov E.V., Shapeev V.P. On the Efficiency of Combining Different Methods for Acceleration of Iterations at the Solution of PDEs by the Method of Collocations and Least Residuals. *Applied Mathematics and Computation*, 2019, vol. 363, pp. 1–19. DOI: 10.1016/j.amc.2019.124644
17. Shapeev V.P., Vorozhtsov E.V. High-Accuracy Numerical Solution of the Second-Kind Integral Equations in the Mathematica Environment. *Journal of Multidisciplinary Engineering Science and Technology*, 2018, vol. 5, no. 12, pp. 9308–9319.
18. Dyakonov V.P. *Mathematica 5/6/7. Polnoe rukovodstvo* [Mathematica 5/6/7. Complete Guide]. Moscow, DMK Press, 2010. (in Russian)

Received July 5, 2023

**РЕШЕНИЕ ЗАДАЧИ КОШИ ДЛЯ ОБЫКНОВЕННЫХ
ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ МЕТОДОМ КОЛЛОКАЦИИ
И НАИМЕНЬШИХ КВАДРАТОВ С АППРОКСИМАЦИЕЙ ПАДЕ***В.П. Шапеев*^{1,2}¹Институт теоретической и прикладной механики им. С.А. Христиановича СО РАН,
г. Новосибирск, Российская Федерация²Новосибирский государственный университет, г. Новосибирск,
Российская Федерация

Предложен и реализован новый способ решения задачи Коши для обыкновенного дифференциального уравнения методом коллокации и наименьших квадратов повышенной точности. Он основан на выводе приближенного нелинейного уравнения многоточечной аппроксимацией рассматриваемой задачи. Приближенное решение задачи в виде аппроксимации Паде сводится к итерационному решению линейной задачи наименьших квадратов относительно коэффициентов искомой дробно-рациональной функции. В случае нелинейных дифференциальных уравнений применяется их предварительная линеаризация. Показано значительное превосходство по точности предложенного в работе способа решения задачи над точностью процедуры NDSolve системы Mathematica. На решении конкретного примера показано превосходство по точности предложенного способа над методом Рунге – Кутты четвертого порядка. Приведены примеры решения предложенным способом задачи Коши для линейных и нелинейных уравнений с точностью, близкой к величине ошибок округлений при операциях на компьютере с числами в формате double. Показано, что точность решения задачи существенно зависит от сложности поведения значений правой части уравнения на заданном отрезке. Приведен пример построения сплайна из кусков аппроксимантов Паде на частичных отрезках, на которые разбит заданный отрезок, в случае, когда требуется повысить точность решения.

Ключевые слова: задача Коши; обыкновенное дифференциальное уравнение; аппроксимация Паде; метод коллокации и наименьших квадратов; повышенная точность; система Mathematica.

Литература

1. Baker, G.A.Jr. Essentials of Pade Approximants / G.A.Jr. Baker. – New York: Academic Press, 1975.
2. Гончар, А.А. О сходимости совместных аппроксимаций Паде для систем функций марковского типа / А.А. Гончар, Е.А. Рахманов // Труды Математического института имени В.А. Стеклова. – 1981. – Т. 157. – С. 31–48.
3. Гончар, А.А. Равновесные распределения и скорость рациональной аппроксимации аналитической функции / А.А. Гончар, Е.А. Рахманов // Математический сборник. – 1987. – Т. 134(176), № 3. – С. 306–352.
4. Гончар, А.А. Многоточечные аппроксимации Паде / А.А. Гончар, Н.Н. Новикова, Г.М. Хенкин // Математический сборник. – 1996. – Т. 187, № 12. – С. 57–86.
5. Суетин, С.П. Аппроксимации Паде и эффективное аналитическое продолжение степенного ряда / С.П. Суетин // Успехи математических наук. – 2002. – Т. 157, № 1. – С. 45–142.
6. Гончар, А.А. Рациональные аппроксимации аналитических функций / А.А. Гончар // Современные проблемы математики. – 2003. – № 1. – С. 83–106.

7. Аптекарев, А.И. Аппроксимации Паде, непрерывные дроби и ортогональные многочлены / А.И. Аптекарев, В.И. Буслаев, А. Мартинес-Финкельштейн, С.П. Суетин // Успехи математических наук. – 2011. – Т. 66, № 6. – С. 37–122.
8. Хованский, А.Н. Приложение цепных дробей и их обобщений к вопросам приближенного анализа. М.: Государственное издательство технико-теоретической литературы, 1956.
9. Величко, И.Г. Применение метода цепных дробей для получения аппроксимаций Паде решений задач Коши для дифференциальных уравнений первого порядка / И.Г. Величко, И.Г. Ткаченко, В.В. Балабанова // Вестник науки и образования Северо-Запада России. – 2015. – Т. 1, № 3. – С. 1–9.
10. Вишневский, В.Э. Аппроксимация Паде решения задачи Коши / В.Э. Вишневский, А.В. Зубов, О.А. Иванова // Вестник Санкт-Петербургского университета. Прикладная математика, серия 10. – 2012. – № 4. – С. 3–17.
11. Исаев, В.И. Варианты метода коллокаций и наименьших квадратов повышенной точности для численного решения уравнений Навье-Стокса / В.И. Исаев, В.П. Шапеев // Журнал вычислительной математики и математической физики. – 2010. – Т. 50, № 10. – С. 1758–1770.
12. Shapeev, V.P. CAS Application to the Construction of the Collocations and Least Residuals Method for the Solution of the Burgers and Korteweg-de Vries–Burgers Equations / V.P. Shapeev, E.V. Vorozhtsov // International Workshop on Computer Algebra in Scientific Computing. Warsaw, 2014. – V. 8660. – P. 432–446.
13. Шапеев, В.П. Варианты метода коллокации и наименьших невязок повышенной точности в области с криволинейной границей / В.П. Шапеев, В.А. Беляев // Вычислительные технологии. – 2016. – Т. 21, № 5. – С. 95–110.
14. Belyaev, V.A. Versions of the Collocation and Least Squares Method for Solving Biharmonic Equations in Non-Canonical Domains / V.A. Belyaev, V.P. Shapeev // AIP Conference Proceedings. – 2017. – V. 1893, № 1. – Article ID: 030102.
15. Shapeev, V.P. New Possibilities and Applications of the Least Squares Collocation Method / V.P. Shapeev, V.A. Belyaev, S.K. Golushko, S.V. Idimeshev // EPJ Web of Conferences. – 2018. – V. 173. – Article ID: 01012.
16. Vorozhtsov, E.V. On the Efficiency of Combining Different Methods for Acceleration of Iterations at the Solution of PDEs by the Method of Collocations and Least Residuals / E.V. Vorozhtsov, V.P. Shapeev // Applied Mathematics and Computation. – 2019. – V. 363. – P. 1–19.
17. Shapeev, V.P. High-Accuracy Numerical Solution of the Second-Kind Integral Equations in the Mathematica Environment / V.P. Shapeev, E.V. Vorozhtsov // Journal of Multidisciplinary. Engineering Science and Technology. – 2018. – V. 5, № 12. – P. 9308–9319.
18. Дьяконов, В.П. Mathematica 5/6/7. Полное руководство / В.П. Дьяконов. – М.: ДМК Пресс, 2010.

Василий Павлович Шапеев, доктор физико-математических наук, главный научный сотрудник, Институт теоретической и прикладной механики им. С.А. Христиановича СО РАН (г. Новосибирск, Российская Федерация); профессор, кафедра «Математическое моделирование», Новосибирский государственный университет (г. Новосибирск, Российская Федерация), shapeev.vasily@mail.ru.

Поступила в редакцию 5 июля 2023 г.