Computational modeling and simulation

On the properties of numerical solutions of dynamical systems obtained using the midpoint method

Vladimir P. Gerdt[1], **Mikhail D. Malykh**[2],
Leonid A. Sevastianov[1,2], **Yu Ying**[2,3]

[1] *Joint Institute for Nuclear Research*
Joliot-Curie St. 6, Dubna, Moscow Region 141980, Russian Federation
[2] *Department of Applied Probability and Informatics*
Peoples' Friendship University of Russia (RUDN University)
Miklukho-Maklaya St. 6, Moscow 117198, Russian Federation
[3] *Department of Algebra and Geometry*
Kaili University
Kaiyuan Road 3, Kaili 556011, China

The article considers the midpoint scheme as a finite-difference scheme for a dynamical system of the form $\dot{x} = f(x)$. This scheme is remarkable because according to Cooper's theorem, it preserves all quadratic integrals of motion, moreover, it is the simplest scheme among symplectic Runge-Kutta schemes possessing this property.

The properties of approximate solutions were studied in the framework of numerical experiments with linear and nonlinear oscillators, as well as with a system of several coupled oscillators. It is shown that in addition to the conservation of all integrals of motion, approximate solutions inherit the periodicity of motion. At the same time, attention is paid to the discussion of introducing the concept of periodicity of an approximate solution found by the difference scheme.

In the case of a nonlinear oscillator, each step requires solving a system of nonlinear algebraic equations. The issues of organizing computations using such schemes are discussed. Comparison with other schemes, including those symmetric with respect to permutation of x and $\hat{x}$.

Key words and phrases: conservative finite-difference schemes, dynamical systems, Sage, Maple

Introduction

One of the most common and popular mathematical models is the Cauchy problem for an autonomous system of ordinary differential equations. Analytical methods make it possible to find the algebraic integrals of motion [1] for such systems, and numerical methods allow approximate plotting the particular solutions [2].

Consider an autonomous system of differential equations in an affine space of dimension n

$$\dot{x} = f(x). \tag{1}$$

Here $x = (x_1, \ldots, x_n)$ is a point in the affine space, $f = (f_1, \ldots, f_n)$ is a set of rational functions belonging to $\mathbb{Q}(x)$. Except the cases where it leads to ambiguity, we will use the vector notation $\dot{x} = f(x)$. By the algebraic integral of motion of this system, we mean the algebraic function g, constant on any particular solution of the system (1), i.e., satisfying the equation

$$f_1 \frac{\partial g}{\partial x_1} + \cdots + f_n \frac{\partial g}{\partial x_n} = 0. \tag{2}$$

It can be shown that the existence of an algebraic integral implies the existence of a rational integral, therefore only rational integrals are considered below [3].

The finite difference method is a standard numerical method for solving systems of ordinary differential equations [2]. The finite-difference scheme for solving the system of equations (1) describes the transition from the value of x taken at some instant of time t to the value of x taken at the next instant of time $t + \Delta t$. This new value will be denoted below by $\hat{x}$. Of course, by the difference scheme for the system (1) we understand a correspondence in some sense approximating the system of differential equations, rather an arbitrary correspondence between the variables x and $\hat{x}$. Usually, by this approximation we mean that the system of equations defining the difference scheme tends to the original system as $\Delta t \to 0$.

Definition 1. By a particular solution of the system (1), found using the finite-difference scheme, we mean a finite or infinite sequence of points

$$x^{(0)}, \ x^{(1)}, \ x^{(2)}, \ldots, x^{(m)}, \ldots$$

of the n-dimensional space (or subset considered in it), the first element of which is taken arbitrarily, and each next element is obtained from the previous one according to the difference scheme:

$$x^{(m+1)} = \hat{x}^{(m)}, \quad m = 0, 1, 2, \ldots$$

This approximate particular solution will be associated with the exact solution of the Cauchy problem

$$\dot{x} = f(x), \quad x|_{t=0} = x^{(0)}.$$

Analytical and numerical methods cannot always be reconciled. It often turns out that the algebraic integral of motion is known, but a difference

scheme is used, which does not preserve this integral. Therefore, in numerical experiments, it is often nothing to do but to observe with regret how the quantity which remains constant on the exact solution, changes step by step. Very frequently, the integrals of motion express fundamental laws of nature, the violation of which introduces new properties into the mathematical model under consideration, or trivial geometric relationships, the violation of which makes it difficult to interpret the results of integration.

The idea of contructing finite-difference schemes exactly preserving the integral of motion of dynamical systems was proposed in late 1980s in the papers by Yu. B. Suris [4] and Cooper [5], approaching the problem from different sides, namely, Yu. B. Suris from composing difference schemes for Hamiltonian systems that preserve a symplectic structure, and Cooper from preserving integrals. As a result, a large family of Runge-Kutta schemes was discovered that preserve all quadratic integrals of any dynamical system and the symplectic structure of the Hamiltonian system [6].

The simplest of this class of schemes is the midpoint scheme. By construction, the approximate solutions found using this scheme retain all quadratic integrals. Traditionally, the question of 'improvement' of convergence due to the conservation of integrals remained in the focus of attention. In this article, we intend to clarify what other qualitative properties of the exact solution are inherited by the approximate one. For completeness, we give an elementary proof of Cooper's theorem.

1.   Conservation of quadratic integrals

Let g be an integral of the system of Eqs. 1. According to Lagrange theorem about the mean value,

$$g(\hat{x}_1, ..., \hat{x}_n) - g(x_1, ..., x_n) = \sum_{i=1}^{n} \frac{\partial g}{\partial x_i}(c_1, ... c_n) \cdot \Delta x_i,$$

where the derivatives are calculated at the point c, lying somewhere in the segment connecting the points x and $\hat{x}$, i.e., $c_i = x_i + \theta \Delta x_i$, $\theta \in (0, 1)$.

The Lagrange theorem ensures the existence of a suitable function $\theta(x, \hat{x})$, taking the values between 0 and 1 at the real values of $x, \hat{x}$. However, it provides no method to find this function. If we knew such a function, we could easily construct a difference scheme that preserves the manifold $g = 0$ exactly. Indeed, let

$$\frac{\Delta x_i}{\Delta t} = f_i \left(x_i + \theta \Delta x_i \right), \quad i = 1, 2, ..., n. \tag{3}$$

Then

$$g(\hat{x}_1, ..., \hat{x}_n) - g(x_1, ..., x_n) = \sum_{i=1}^{n} \frac{\partial g}{\partial x_i}(c_1, ... c_n) \cdot f_i(c_1, ..., c_n),$$

and this expression is exactly equal to zero due to (2).

For some classes of functions, the Lagrange theorem allows a constructive formulation.

Lemma 1. *If g is a polynomial from $\mathbb{C}[x_1, \dots x_n]$, the degree of which does not exceed 2 inclusively, then*

$$g(\hat{x}_1, \dots, \hat{x}_n) - g(x_1, \dots, x_n) = \sum_{i=1}^{n} \frac{\partial g}{\partial x_i}(c_1, \dots c_n) \cdot \Delta x_i,$$

where $c_i = (\hat{x}_i + x_i)/2$.

Proof. Let u be a new auxiliary variable and

$$G(u) = g(x_1 + u\Delta x_1, \dots, x_n + u\Delta x_n).$$

According to the Lagrange theorem, there is such value $\theta \in (0, 1)$, that

$$G(1) - G(0) = G'(\theta), \tag{4}$$

or, in more detail,

$$g(\hat{x}_1, \dots, \hat{x}_n) - g(x_1, \dots, x_n) = \sum_{i=1}^{n} \frac{\partial g}{\partial x_i}(c_1, \dots c_n) \cdot \Delta x_i,$$

where $c_i = x_i + \theta \Delta x_i$.

We have to prove that $\theta = 1/2$.

By the hypothesis of the lemma, g is a polynomial whose degree does not exceed 2, therefore G'' is a polynomial with respect to u whose degree does not exceed 2, i.e.

$$G(u) = au^2 + bu + c, \quad G'(u) = 2au + b.$$

But then the Lagrange formula (4) reduces to $a + b = 2a\theta + b$, from which it immediately follows that $\theta = 1/2$. □

The finite-difference scheme (3) at $\theta = 1/2$, i.e., the scheme

$$\frac{\Delta x_i}{\Delta t} = f_i\left(\frac{\hat{x}_1 + x_1}{2}, \dots, \frac{\hat{x}_n + x_n}{2}\right), \quad i = 1, 2, \dots, n, \tag{5}$$

is called the midpoint scheme. Lemma 1 immediately leads to the following theorem.

Theorem 1 (Cooper, 1983; [5]). *The midpoint finite-difference scheme (5) not only approximates the system (1), but also preserves all linear and quadratic integrals of this system.*

2. Harmonic oscillator

Consider the simplest dynamical system

$$\begin{cases} \dot{x} = -y, \\ \dot{y} = x, \end{cases} \tag{6}$$

describing a harmonic oscillator. This system is Hamiltonian, the energy conservation law for it yields the integral $x^2 + y^2 = C$.

Usually it is immediately said without further ado that the solutions of this system describe periodic rotations along concentric circles in the phase plane xy.

The standard discretization does not allow to preserve this description. Consider for definiteness, the explicit Euler scheme

$$\begin{cases} \hat{x} - x = -y\Delta t, \\ \hat{y} - y = x\Delta t. \end{cases}$$

Figure 1 presents the curve in the phase plane, obtained using the Euler method instead of a unit circle. Thus it is easy to describe the difference between the exact and approximate solutions. Instead of closed curves, a spiral in the phase plane appears, or, in terms of classical mechanics, the time discretization leads to gradual increase of the system energy.
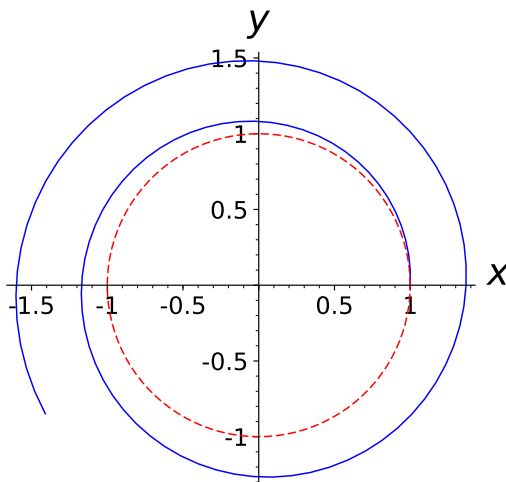

Figure 1. Solution of the initial value problem for Eqs. (6) using the Euler method (solid) and the meanpoint method (dashed), $\Delta t = 0.1$, 100 steps are maid

Now let us consider the meanpoint scheme

$$\hat{x} - x = -(\hat{y} + y)\frac{\Delta t}{2}, \quad \hat{y} - y = (\hat{x} + x)\frac{\Delta t}{2} \tag{7}$$

and try to understand the qualitative difference between the approximate solution and the exact one. According to Cooper's theorem, this difference scheme exaclty preserves this integral. Therefore, in the phase plane we get ovals that seemingly do not differ from a circle (see Figure 1). However, these curves are still not closed, and the motion along them cannot be considered perodic, since the values of x, y do not repeat. Therefore, it could be executed that the absence of closedness and periodicity completely allows one to distinguish between the approximate solution and the exact one. In

fact, it is possible to achieve the solution relativity by the appropriate step choice.

Theorem 2. *If we take for α the minimal positive root of the equation*

$$(1 + i\alpha)^{2N} = (1 + \alpha^2)^N, \tag{8}$$

which in terms of trigonometric functions can be expressed as

$$\alpha = \tan \frac{\pi}{N},$$

then the calculation according to the meanpoint finite-difference scheme (7) *with the step*

$$\Delta t = 2\alpha = 2 \tan \frac{\pi}{N},$$

in N steps leads to the initial values of x, y.

Remark 1. For us it is convenient to use the transcendent formula $\alpha = \tan \frac{\pi}{N}$, however, one could do without it using purely algebraic means. To emphasize this fact, in the statement of the theorem we have preserved the algebraic equation, for which this number is a root.

For proof, let us express $\hat{x}, \hat{y}$ in terms of x, y, denoting for brevity

$$\frac{\Delta t}{2} = \alpha.$$

Then

$$\begin{cases} \hat{x} + \alpha\hat{y} = x - \alpha y, \\ -\alpha\hat{x} + \hat{y} = \alpha x + y \end{cases}$$

or, in the matrix form

$$\begin{pmatrix} 1 & \alpha \\ -\alpha & 1 \end{pmatrix} \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} = \begin{pmatrix} 1 & -\alpha \\ \alpha & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Inverting the matrix, we get

$$\begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} = \frac{1}{1 + \alpha^2} \begin{pmatrix} 1 & -\alpha \\ \alpha & 1 \end{pmatrix}^2 \begin{pmatrix} x \\ y \end{pmatrix} \tag{9}$$

now the proof of theorem 2 is reduced to the proof of the following purely algebraic lemma.

Lemma 2. *If we take for α the minimal positive root of equation* (8), *which in terms of trigonometric functions can be written as*

$$\alpha = \tan \frac{\pi}{N},$$

then A^{2N} is a unit matrix.

Proof. Set for brevity

$$A = \frac{1}{\sqrt{1+\alpha^2}} \begin{pmatrix} 1 & -\alpha \\ \alpha & 1 \end{pmatrix}.$$

Assume that we performed N steps of the finite-difference scheme, having started from the values x, y at a certain value $t = 0$, and finished with x^N, y^N at $t = N\Delta t$. Then

$$\begin{pmatrix} x^N \\ y^N \end{pmatrix} = A^2 \begin{pmatrix} x^{N-1} \\ y^{N-1} \end{pmatrix} = \cdots = A^{2N} \begin{pmatrix} x \\ y \end{pmatrix}.$$

The eigenvalues of the matrix A are

$$\lambda_1 = \frac{1+i\alpha}{\sqrt{1+\alpha^2}}, \quad \lambda_2 = \frac{1-i\alpha}{\sqrt{1+\alpha^2}}.$$

When a matrix is raised to a power, its eigenvalues are raised to this power, too.

The equation

$$\lambda_1^{2N} = 1 \tag{10}$$

can be satisfied assuming that

$$\frac{1+i\alpha}{\sqrt{1+\alpha^2}} = \cos\frac{\pi}{N} + i\sin\frac{\pi}{N}$$

or

$$\alpha = \tan\frac{\pi}{N}.$$

The equality (10) can be rewritten in the form of the algebraic equation (8); in this way the minimal positive root of this equation has been found.

At real Δt from the equality (10) the equality $\lambda_2^{2N} = 1$ follows. In this case both eigenvalues of the matrix A^{2N} are equal to 1. The eigenvectors of matrix A are also eigenvectors of any power of this matrix, therefore, in this case two linearly independent eigenvectors correspond to the unit eigenvalue. This is possible only if the matrix A^{2N} coincides with the unit matrix.   □

The proved theorem gives rise to the following definition.

Definition 2. The particular solution

$$x, x', x^{(2)}, \ldots$$

found using a certain finite-difference scheme approximating the system (1), at a certain numerical value of the step Δt will be called periodic, if for some $N \in \mathbb{N}$

$$x^{(N)} = x.$$

The number $N\Delta t$ will be called a period of this particular solution.

Theorem 2 means that the midpoint scheme (7) yields periodic solutions at a number of step values that form a descending sequence:

$$\Delta t_3 = 2\tan\frac{\pi}{3} = 2\sqrt{3} = 3.46,$$

$$\Delta t_4 = 2\tan\frac{\pi}{4} = 2,$$

$$\Delta t_5 = 2\tan\frac{\pi}{5} = 2\sqrt{-2\sqrt{5}+5} = 1.45,$$

$$\Delta t_6 = 2\tan\frac{\pi}{6} = \frac{2}{3}\sqrt{3} = 1.15,$$

...

$$\Delta t_{20} = 2\tan\frac{\pi}{20} = 2\sqrt{5} - 2\sqrt{2\sqrt{5}+5} + 2 = 0.31,$$

...

converging to zero. The corresponding sequence of periods

$$N\Delta t_N = 2N\tan\frac{\pi}{N} = 2\pi + \frac{2\pi^3}{3}\frac{1}{N^2} + ...$$

converges to the exact solution period 2π.

Thus, the second most important qualitative property of the exact solution, i.e., its periodicity, can be preserved, too. Moreover, even for small values of N, in this way it is possible to get a solution that is qualitatively similar to the exact one.

For example, for $N = 10$ and the step value

$$\Delta t = 2\tan\frac{\pi}{10} = \frac{2}{5}\sqrt{-10\sqrt{5}+25} = 0.64$$

we arrive at a periodic scheme with the period $T = 10\Delta t$, that differs from the period of the exact solution by a noticeable value of

$$T - 2\pi = 0.21.$$

In the phase plane xy a regular decagon is obtained that obviously differs from a circle, too. However, even in this case the trajectory in the phase plane is exactly closed and the motion is periodic. In other words, the approximate solution found using the midpoint scheme turns out to be qualitatively similar to the exact one, being rather different quantitatively.

In our opinion, when using the harmonic oscillator model, there is not enough reason to think that the characteristic time scale, taken as physically small, is really small from the point of view of computational mathematics. Therefore, the real reason in favor of using the continuous form of the Newton equations rather than the discrete one is that the traditional method of discretization leads to a violation of the fundamental laws of mechanics (the law of energy conservation) and to a solution, whose properties differ from the expected periodicity. The use of periodic difference schemes removes this difficulty.

3.   A system of coupled oscillators

A system of coupled oscillators can be described as a Hamiltonian system with the Hamiltonian

$$H = \sum_{i=1}^{n}\sum_{j=1}^{n} m_{ij}y_i y_j - \sum_{i=1}^{n}\sum_{j=1}^{n} k_{ij}x_i x_j.$$

It is assumed that the kinetic and potential energy

$$\sum_{i=1}^{n} m_{ij}y_i y_j, \quad \sum_{i=1}^{n} k_{ij}x_i x_j$$

are described by positively defined quadratic forms.

In the matrix form the system is written as $M\ddot{\vec{x}} = -K\vec{x}$.

Since the matrices K and M are symmetric and positively defined, the generalized eigenvalue problem $K\vec{\xi} = \lambda M\vec{\xi}$ yields n eigenvectors $\vec{\xi}_1, \ldots, \vec{\xi}_n$ that form a basis $\mathbb{R}^n$, orthonormalized with the weight M, corresponding to n positive eigenvalues $\lambda_1 = \omega_1^2, \ldots, \lambda_n = \omega_n^2$. In this basis

$$\vec{x} = \sum_{i=1}^{n} z_i(t)\vec{\xi}_i,$$

where $z_i = \vec{\xi}_i^T M\vec{x}$.

The differential equation in these variables is separated into n independent equations

$$\vec{\xi}_j^T M \sum_{i=1}^{n} \ddot{z}_i \vec{\xi}_i = \vec{\xi}_j^T K \sum_{i=1}^{n} z_i \vec{\xi}_i$$

or, due to the basis orthogonality, $\ddot{z}_j + \omega_j^2 z_j = 0$, $j = 1, \ldots, n$.

Each of these equations separately describes the vibration of a harmonic oscillator with the frequency ω_j. Therefore, the positive numbers $\omega_1, \ldots, \omega_n$ are called eigenfrequencies of the system of coupled oscillators. Now from the formula

$$\vec{x} = \sum_{i=1}^{n} z_i(t)\vec{\xi}_i$$

it is seen that the oscillations of the coupled system will be a superposition of harmonic oscillations at the eigenfrequencies.

Now it is important for us that this system has n algebraic integrals

$$\dot{z}_j^2 + \omega_j^2 z_j^2 = C_j,$$

which in old variables can be written as

$$(\vec{\xi}_i^T M\dot{\vec{x}})^2 + \omega_j^2 (\vec{\xi}_i^T M\vec{x})^2 = C_j, \quad j = 1, 2, \ldots, n. \tag{11}$$

Thus, the initial system has n quadratic integrals. These integrals will be called partial integrals.

Lemma 3. *If the eigenfrequencies are incommensurable, then any algebraic integral of a system of coupled oscillators is expressed algebraically in terms of n partial integrals.*

The proof of this lemma is somewhat lengthy and is not given here. In essence, it relies on constructions from the proof of Bruns theorem proposed by Painlevé [7].

Remark 2. The condition of incommensurability of frequencies is essential. If $\omega_1 = \omega_2$, then the expression

$$\dot{z}_1 z_2 - z_1 \dot{z}_2 = C$$

is an additional algebraic integral. In fact, the general solution of the system is

$$z_1 = C_1 \cos \omega_1 t + C_2 \sin \omega_1 t, \quad z_2 = C_3 \cos \omega_1 t + C_4 \sin \omega_1 t, ...$$

therefore $\dot{z}_1 z_2 - z_1 \dot{z}_2 = (C_2 C_3 - C_1 C_4)\omega_1$, i.e., is really a constant. Thus, without a separate study, it cannot be ruled out that in degenerate cases the system has additional integrals of motion that are not preserved by the midpoint scheme.

According to the theorem 1, the midpoint scheme preserves all partial integrals, and, therefore, all algebraic integrals of the system.

The step Δt can always be chosen so that the normal oscillation at a certain natural frequency is described by a periodic particular solution. However, this step depends on the harmonic number. Therefore, it is impossible to choose a time step at which the oscillation, in the framework of the continuous model, which is a superposition of periodic normal oscillations (mixed oscillation), would be the sum of periodic particular solutions.

Thus, for the case of linear systems of ordinary differential equations, the midpoint scheme is a difference scheme that preserves all algebraic integrals. Moreover, for periodic particular solutions, it is possible to select a step in a purely algebraic way so that the approximate solution becomes periodic in the sense of the definition 2.

4.   Elliptic oscillator

We now proceed to the nonlinear case. Solid state dynamics provides many excellent examples of autonomous systems with periodic solutions. The simplest of them are integrated in Jacobi elliptic functions [8].

By the definition of Jacobi functions [9],

$$p = \operatorname{sn} t, \quad q = \operatorname{cn} t, \quad r = \operatorname{dn} t$$

is a particular solution of the nonlinear autonomous system of differential equations

$$\dot{p} = qr, \ \dot{q} = -pr, \ \dot{r} = -k^2 pq \tag{12}$$

with the initial conditions $p = 0$, $q = r = 1$ at $t = 0$.

This autonomous system has two quadratic integrals of motion

$$p^2 + q^2 = \text{const} \quad \text{and} \quad k^2 p^2 + r^2 = \text{const.} \tag{13}$$

This system is convenient for us, because its properties are not only well studied, but the solutions themselves are easily accessible in any computer algebra system, e.g., in Sage [10].

The system (12) is a convenient test for studying the conservatism of finite-difference schemes used in the common numerical solvers of ordinary differential equations. We, in cooperation with Yu.A. Blinkov, conducted tests on the following systems:

1. lsoda: Real-valued Variable-coefficient Ordinary Differential Equation solver, with fixed-leading-coefficient implementation. It provides automatic method switching between implicit Adams method (for non-stiff problems) and a method based on backward differentiation formulas (BDF) (for stiff problems). Source: `http://www.netlib.org/odepack`;

2. vode: Real-valued Variable-coefficient Ordinary Differential Equation solver, with fixed-leading-coefficient implementation. It provides implicit Adams method (for non-stiff problems) and a method based on backward differentiation formulas (BDF) (for stiff problems). Source: `http://www. netlib.org/ode/vode.f`;

3. dopri5, dop853: This is an explicit Runge-Kutta method of order (4)5 due to Dormand & Prince (with stepsize control and dense output).

In Figures 4–7 it is well seen that in all cases the value of $p^2 + q^2$ increases or decreases almost linearly. Only in the first solver (see Figures 2, 3) "random" fluctuations are observed, but with a trend towards linear growth.
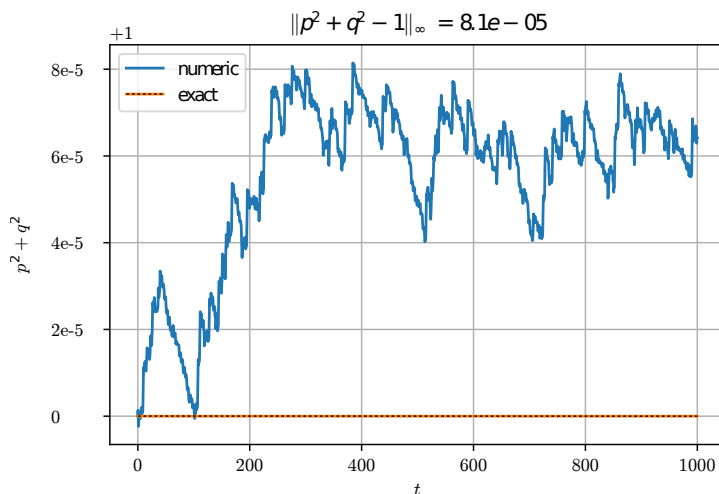

Figure 2. The value of $p^2 + q^2 - 1$, calculated with the solver Lsoda using the method based on backward differentiation formulas

From an analytical point of view, this system is remarkable because any particular solution of it is representable as the ratio of two everywhere convergent series in powers of t [8]. From the point of view of the finite difference method, this system is remarkable because it can be approximated by a difference scheme, namely, the midpoint scheme (5), which preserves its integrals exactly.
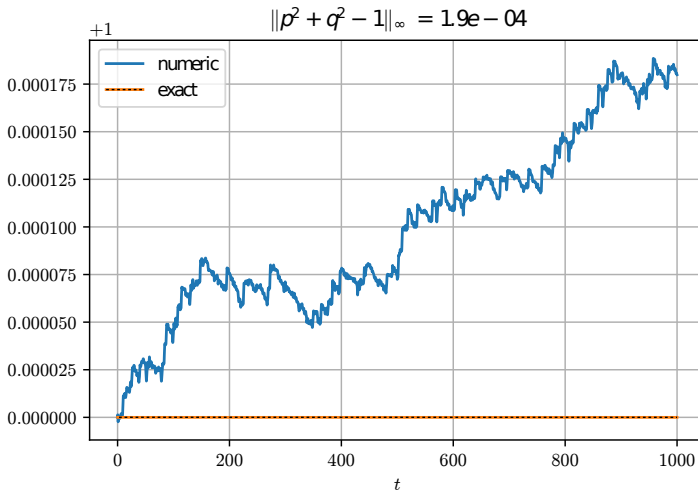
Figure 3. The value of $p^2 + q^2 - 1$, calculated with the solver Lsoda using the method based on the default settings
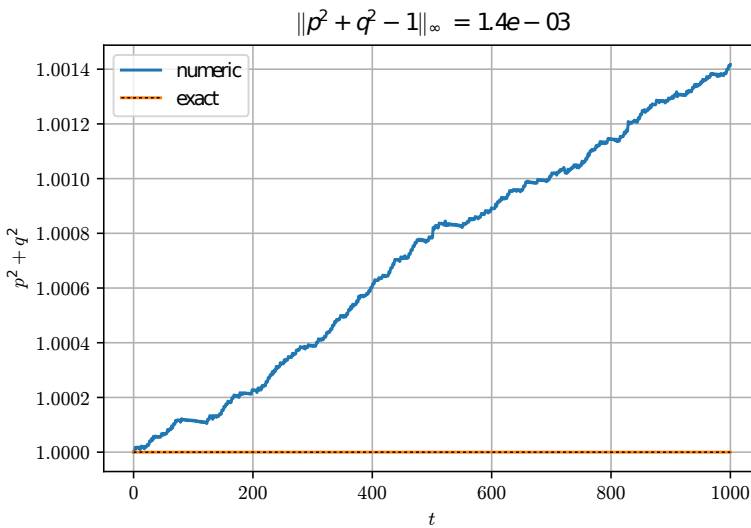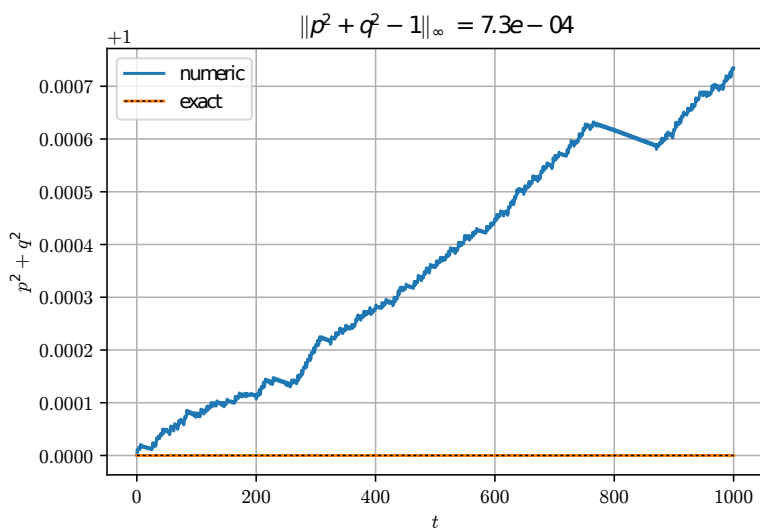

Figure 4. The value of $p^2 + q^2 - 1$, calculated with the solver Vode using the method based on backward differentiation formulas (BDF) with the choice of Jacobian

However, for nonlinear differential equations, the organization of the transition from layer to layer according to the midpoint scheme requires solving a system of nonlinear algebraic equations. The organization of calculations according to such a scheme is a subject of discussion.

Figure 5. The value of $p^2 + q^2 - 1$, calculated with the solver Vode using the method based on backward differentiation formulas (BDF) without the choice of Jacobian
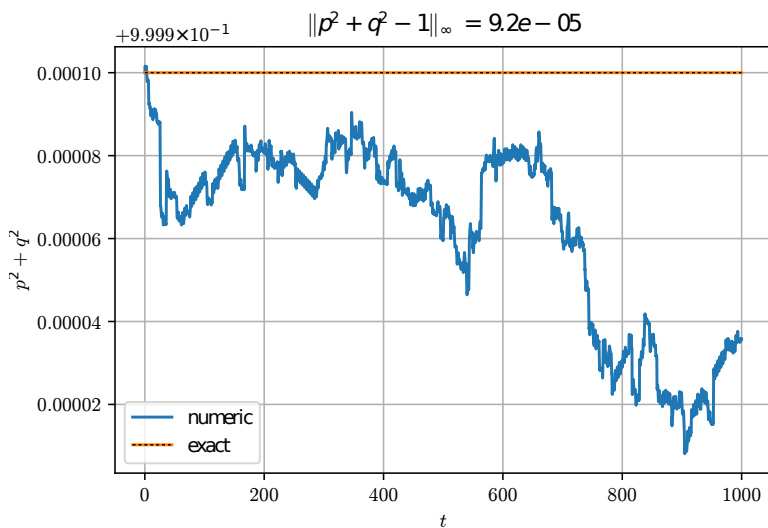

Figure 6. The value of $p^2 + q^2 - 1$, calculated with the solver Vode using the default settings

The midpoint scheme for the system (12) is written as follows:

$$\begin{cases} \dfrac{\hat{p} - p}{\Delta t} - \dfrac{\hat{q} + q}{2}\dfrac{\hat{r} + r}{2} = 0, \\[3mm] \dfrac{\hat{q} - q}{\Delta t} + \dfrac{\hat{p} + p}{2}\dfrac{\hat{r} + r}{2} = 0, \\[3mm] \dfrac{\hat{r} - r}{\Delta t} + k^2\dfrac{\hat{p} + p}{2}\dfrac{\hat{q} + q}{2} = 0. \end{cases} \qquad (14)$$
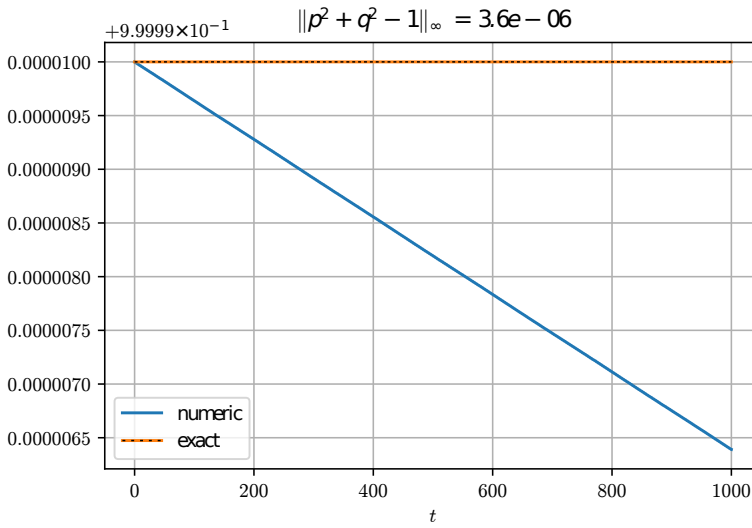
Figure 7. The value of $p^2 + q^2 - 1$, calculated with the solver Dopri5

To pass from one layer to another it is necessary from the given numerical values of Δt and p, q, r to find $\hat{p}$, $\hat{q}$, $\hat{r}$, having solved the system of nonlinear equations. In this case, in addition to the required root, close to the values of p, q, r at small Δt, extraneous roots are also obtained.

Let us investigate this system in Sage, replacing the hat notation with doubling the letter, e.g., using the notation pp instead of $\hat{p}$. The system of equations (14) generates an ideal J in the ring $\mathbb{Q}[p, q, r, \hat{p}, \hat{q}, \hat{r}, \Delta t, k]$.

In Sage this ideal can be specified by the following code:

```
sage: vars=var('p,q,r,pp,qq,rr,dt,k')
sage: K=QQ[vars]
sage: eqs=[4*(pp-p)-(q+qq)*(r+rr)*dt,\
4*(qq-q)+(p+pp)*(r+rr)*dt,\
4*( rr-r)+k^2*(p+pp)*(q+qq)*dt]
sage: J=K.ideal(eqs)
```

Now we can compose an equation connecting $\hat{p}$ with p, q, r by means of elimination ideals [11]:

```
sage: J.elimination_ideal([K(qq),K(rr)]).gens()[0]
```

as

$$p^5 dt^4 k^4 + 3p^4 \hat{p} dt^4 k^4 + 2p^3 \hat{p}^2 dt^4 k^4 - 2p^2 \hat{p}^3 dt^4 k^4 - 3p\hat{p}^4 dt^4 k^4 - \hat{p}^5 dt^4 k^4$$

$$+16p^2 qr dt^3 k^2 + 32pqr\hat{p} dt^3 k^2 + 16qr\hat{p}^2 dt^3 k^2 - 32p^3 dt^2 k^2 - 64pq^2 dt^2 k^2$$

$$-32p^2 \hat{p} dt^2 k^2 - 64q^2 \hat{p} dt^2 k^2 + 32p\hat{p}^2 dt^2 k^2 + 32\hat{p}^3 dt^2 k^2 - 64pr^2 dt^2 - 64r^2 \hat{p} dt^2$$

$$+256qr dt + 256p - 256\hat{p} = 0.$$

Thus, one value of p, q, r on the previous layer generally corresponds to 5 different values of $\hat{p}$ rather than one value. Since the higher degrees of $\hat{p}$ enter this equation with the factor Δt, for $\Delta t \to 0$ only one of these roots tends

to a finite value, which is obviously $\hat{p} = p$, and the four others go to infinity. Thus, applying the midpoint scheme to a nonlinear system requires solving a nonlinear equation and then choosing from its roots one root that is 'close' to the values in the previous step.

Traditionally, such equations are solved numerically by iterative methods, and the values of p, q, r are used as the first approximation for $\hat{p}$, $\hat{q}$, $\hat{r}$. If the step Δt is sufficiently small, we can expect the fast convergence of the described method. At the same time, it is not possible to control the error of the iterative method due to the extreme cumbersomeness of the known estimates, and instead the number of iterations is simply fixed. However, as correctly noted in Numerical Recipies [12], there are no universal numerical methods for solving systems of algebraic equations.

Symbolic methods, primarily the Gröbner basis technique, allow us to propose a different approach, circumvent this difficulty, and implement calculations according to the midpoint scheme in a different way:

— the first stage (symbolic analysis of the finite-difference scheme): find equations having the form $P(\hat{p}, p, q, r, \Delta t) = 0$, $Q(\hat{q}, p, q, r, \Delta t) = 0$, $R(\hat{r}, p, q, r, \Delta t) = 0$, that follow from Eq. (14); the first of these equations was written out above;

— the second stage (computations with floating point): to pass from layer to layer solve three uncoupled equations to find $\hat{p}, \hat{q}, \hat{r}$ and select the roots close to the values of p, q, r at the previous layer.

Thus, from the numerical solution of the system of nonlinear equations, we proceed to the solution of one algebraic equation of the 5th degree with numerical coefficients. The methods of numerical solution are well established and in practice give errors close to the errors in calculating a radical from a number.

Remark 3. Unfortunately, the algorithms implemented in Sage do not give us, together with the roots, an estimate of this error. The question of the precision solution of algebraic equations with real coefficients is currently discussed at all conferences on numerical methods and computer algebra (MMA'2018, CASC'2019). We hope that in the near future this issue will be completely closed.

We implemented both approaches to the implementation of the midpoint method and made sure that the first approach gives the same result faster. Figures 8–9 show the elliptic sine plot calculated in Sage using the built-in algorithm for calculating elliptic functions and directly using the midpoint method. It is clearly seen that even at extremely large values of t, the oscillation amplitude does not decrease. Thus, when using the midpoint scheme:

— both algebraic integrals of motion are preserved exactly,
— rounding error does not accumulate in a way that is noticeable in numerical experiments,
— the periodic nature of the motion is preserved.

A theoretical study of the last two issues faces significant difficulties, which, in our opinion, make us search not for any conservative schemes, but for schemes in which the transition from layer to layer requires solving linear equations. One such scheme for the elliptic oscillator (12) was specified by us

in [13]; in the general case of the system (1) there are fundamental obstacles to the existence of such schemes [14].
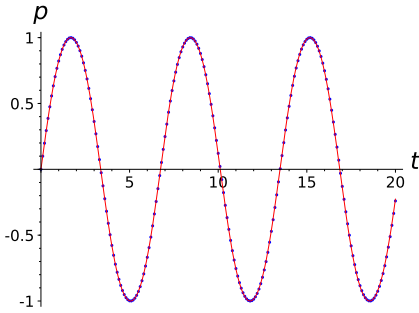

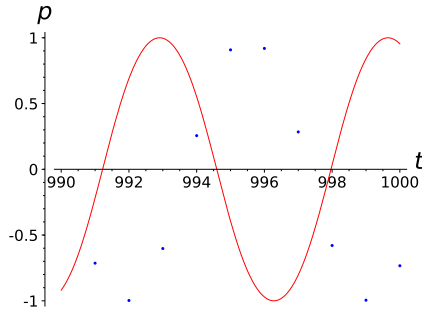Figure 8. sn $\left(t, \frac{1}{2}\right)$, $N = 200$

Figure 9. sn $\left(t, \frac{1}{2}\right)$, $N = 200$

5.   Comparing midpoint scheme with other symmetric schemes

One could think that the conservation of integrals is the result of the scheme symmetry. Therefore, for completeness, we compare the midpoint scheme with another popular symmetric second-order scheme:

$$\frac{\Delta x}{\Delta t} = \frac{f(\hat{x}) + f(x)}{2}. \tag{15}$$

The results of our numerical experiments, presented in the Figure 10, show that this scheme does not preserve the integrals, but gives them values that oscillate with a constant period around some fixed value. This can be accepted for preserving the integrals of the motion on average.
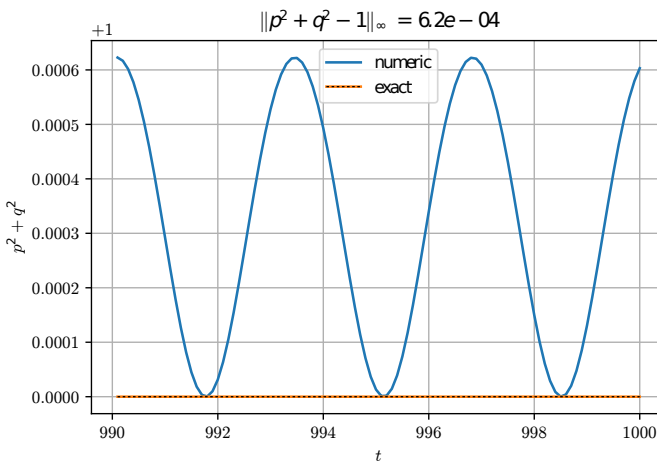

Figure 10. The value of $p^2 + q^2 - 1$, calculated using the scheme (15)

This effect can be explained as follows. The scheme (15) for the Jacobi system yields

$$\Delta p = \frac{\widehat{q}\widehat{r} + qr}{2}\Delta t, \quad \Delta q = -\frac{\widehat{p}\widehat{r} + pr}{2}\Delta t, \quad \Delta r = -k^2\frac{\widehat{p}\widehat{q} + pq}{2}\Delta t. \qquad (16)$$

Therefore

$$\Delta(p^2 + q^2) = (\widehat{p} + p)\Delta p + (\widehat{q} + q)\Delta q =$$

$$= [(\widehat{p} + p)(\widehat{q}\widehat{r} + qr) - (\widehat{q} + q)(\widehat{p}\widehat{r} + pr)]\frac{\Delta t}{2} =$$

$$= [p\widehat{q}\widehat{r} + \widehat{p}qr - \widehat{q}pr - q\widehat{p}\widehat{r}]\frac{\Delta t}{2} =$$

$$= (p\widehat{q} - \widehat{p}q)\Delta r\frac{\Delta t}{2} = (p\Delta q - q\Delta p)\Delta r\frac{\Delta t}{2} =$$

$$= -[p(\widehat{p}\widehat{r} + pr) + q(\widehat{q}\widehat{r} + qr)]\Delta r\frac{\Delta t^2}{4} =$$

$$= [(p\widehat{p} + q\widehat{q})\widehat{r} + (p^2 + q^2)r](\widehat{p}\widehat{q} + pq)\frac{k^2\Delta t^3}{8} \qquad (17)$$

or, expanding in series in powers of Δt

$$\Delta(p^2 + q^2) = (p^2 + q^2)pqr\frac{k^2\Delta t^3}{2} + \dots.$$

For small Δt in the plot of $p^2 + q^2$ versus the step number (or, which is similar, versus t) we will see periodic oscillations of the first term

$$(p^2 + q^2)pqr\frac{k^2\Delta t^3}{2}.$$

Thus, the oscillation of the value of the integral $p^2 + q^2$ near its exact value 1 observed in the numerical experiment is not related to the conservativeness of the scheme, but is due to the fact that the main term in the expansion of the increment of this integral in power of t depends on the values of p, q, r, approximating periodic functions. The considered representation scheme is a perfect example of the scheme symmetric with respect to permutation of x and $\widehat{x}$.

Conclusion

For the study of dynamical systems with quadratic integrals, the midpoint scheme is perfect. This scheme preserves all the integrals of these systems precisely, that is, discretization does not introduce any dissipativity or antidissipativity into the model.

Linear problems have just quadratic integrals, and the calculation according to the midpoint scheme requires solving linear equations, so the issue of constructing conservative difference schemes for linear differential equations

can be considered closed. It worth noting that the Kepler problem after passing to proper time and introducing the Runge-Lenz-Laplace vector turns into a linear dynamical system, therefore, it is possible to construct for it a conservative finite-difference scheme [15].

Conservation of integrals leads to the preservation of a number of qualitative properties of the model, e.g., the periodicity of solutions and the closedness of phase trajectories.

The calculation according to the midpoint scheme for nonlinear systems, even if all their integrals are quadratic, makes it necessary to solve a nonlinear system of algebraic equations at each step, which significantly complicates both the calculation and the study of the properties of approximate solutions. Overcoming this difficulty is the main challenge of the theory of difference schemes for ordinary differential equations.

Acknowledgments

References

[1] A. Goriely, *Integrability and nonintegrability of dynamical systems.* Singapore; River Edge, NJ: World Scientific, 2001.

[2] E. Hairer, G. Wanner, and C. Lubich, *Geometric numerical integration. Structure-preserving algorithms for ordinary differential equations.* Berlin Heidelberg New York: Springer, 2000.

[3] L. Königsberger, *Die Principien der Mechanik.* Leipzig: Teubner, 1901.

[4] Y. B. Suris, "Hamiltonian methods of Runge–Kutta type and their variational interpretation [Gamil'tonovy metody tipa Runge–Kutty i ikh variatsionnaya traktovka]," *Matematicheskoye modelirovaniye*, vol. 2, no. 4, pp. 78–87, 1990, in Russian.

[5] G. J. Cooper, "Stability of Runge–Kutta methods for trajectory problems," *IMA Journal of Numerical Analysis*, vol. 7, pp. 1–13, 1987. DOI: `10.1093/imanum/7.1.1`.

[6] J. M. Sanz-Serna, "Symplectic Runge–Kutta schemes for adjoint equations, automatic differentiation, optimal control, and more," *SIAM REVIEW*, vol. 58, pp. 3–33, 1 2016. DOI: `10.1137/151002769`.

[7] P. Painlevé, "Mémore sur les intégrales du problème des n corps," in *Œuvres de Paul Painlevé.* 1975.

[8]   V. V. Golubev, *Vorlesungen über Differentialgleichungen im Komplexen*. Leipzig: VEB Deutscher Verlag der Wissenschaften, 1958.

[9]   P. F. Byrd and M. D. Friedman, *Handbook of elliptic integrals for engineers and scientists*. Springer, 1971. DOI: `10.1007/978-3-642-65138-0`.

[10] W. A. Stein. (2015). Sage Mathematics Software (Version 6.7), The Sage Development Team, [Online]. Available: `http://www.sagemath.org`.

[11] D. A. Cox, J. Little, and D. O'Shea, *Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra*, 4th ed. Springer, 2015. DOI: `10.1007/978-3-319-16721-3`.

[12] *Numerical Recipes: The Art of Scientific Computing*, 3rd ed. Cambridge University Press, 2007, 1256 pp.

[13] E. A. Ayryan, M. D. Malykh, L. A. Sevastianov, and Yu Ying, "Finite difference schemes and classical transcendental functions," *LNCS*, vol. 11189, pp. 3–33, 2019. DOI: `10.1007/978-3-030-10692-8_26`.

[14] E. A. Ayryan, M. D. Malykh, L. A. Sevastianov, and Yu Ying, "On explicit difference schemes for autonomous systems of differential equations on manifolds," *LNCS*, vol. 11661, pp. 343–361, 2019. DOI: `10.1007/978-3-030-26831-2_23`.

[15] R. Kozlov, "Conservative discretizations of the Kepler motion," *Journal of Physics A*, vol. 40, no. 17, pp. 4529–4539, 2007. DOI: `10.1088/1751-8113/40/17/009`.

[16] Y. A. Blinkov and V. P. Gerdt, "On computer algebra aided numerical solution of ODE by finite difference method," in *Polynomial Computer Algebra '2019. April 15–20, 2019. Euler International Mathematical Institute, St. Petersburg, RUSSIA*. 2019.

Information about the authors:

Vladimir P. Gerdt — Doctor of Physical and Mathematical Sciences, Full Professor at the Joint Institute for Nuclear Research (JINR) where he is the head of the Group of Algebraic and Quantum Computations (e-mail: `gerdt@jinr.ru`, phone: + 7 (49621) 63437, ORCID: https://orcid.org/0000-0002-0825-1811, ResearcherID: T-5179-2019, Scopus Author ID: 57209076002)

Mikhail D. Malykh (Russian Federation) — 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: `malykh-md@rudn.ru`, phone: +7(495)9550927, ORCID: https://orcid.org/0000-0001-6541-6603, ResearcherID: P-8123-2016, Scopus Author ID: 6602318510)

Yu Ying (China)   — postgraduate student of Department of Applied Probability and Informatics of Peoples' Friendship University of Russia (RUDN University); assistant professor of Department of Algebra and Geometry, Kaili University (e-mail: yingy6165@gmail.com, phone: +7(495)9550927,   ORCID: https://orcid.org/0000-0002-4105-2566, ResearcherID: AAC-8344-2019, Scopus Author ID: 57208127921)

Leonid A. Sevastianov (Russian Federation)   — professor, Doctor of Physical and Mathematical Sciences, professor of Department of Applied Probability and Informatics of Peoples' Friendship University of Russia (RUDN University) (e-mail: sevastianov-la@rudn.ru, phone: +7(495)955-07-83,   ORCID: https://orcid.org/0000-0002-1856-4643, ResearcherID: B-8497-2016, Scopus Author ID: 8783969400)

О свойствах численных решений динамических систем, полученных по схеме средней точки

В. П. Гердт[1], М. Д. Малых[2], Л. А. Севастьянов[1,2], Юй Ин[2,3]

[1] *Объединенный институт ядерных исследований*
ул. Жолио Кюри, д. 6, г. Дубна, Московская область, 141980, Россия
[2] *Кафедра прикладной информатики и теории вероятностей*
Российский университет дружбы народов
ул. Миклухо-Маклая, д. 6, г. Москва, 117198, Россия
[3] *Кафедра алгебры и геометрии*
Университет Каили
Kaiyuan Road 3, Каили 556011, Китай

В статье рассматривается схема средней точки как разностная схема для динамической системы вида $\dot{x} = f(x)$. Эта схема замечательна тем, что в силу теоремы Купера сохраняет все квадратичные интегралы движения, более того, это — простейшая схема из числа симплектических схем Рунге–Кутты, обладающих названным свойством.

Свойства приближённых решений изучены в рамках численных экспериментов с линейным и нелинейным осцилляторами, а также с системой нескольких связанных осцилляторов. Показано, что помимо сохранения всех интегралов движения, приближённые решения наследуют периодичность движения. При этом уделено внимание обсуждению введения понятие периодичности приближённого решения, найденного по разностной схеме.

В случае нелинейного осциллятора выполнение каждого шага требует решения системы нелинейных алгебраических уравнений. Обсуждены вопросы организации вычислений по таким схемам. Дано сравнение с другими схемами, в том числе симметрическими относительно перестановки x и $\hat{x}$.

Ключевые слова: консервативные конечно-разностные схемы, динамические системы, Sage, Maple