CHIN.PHYS.LETT. Vol. 25, No. 3 (2008) 866

Structure-Preserving Algorithms for the Lorenz System *

GANG Tie-Qiang(江铁强)[1], MEI Feng-Xiang(梅凤翔)[1], CHEN Li-Jie(陈立杰)[2]
Department of Mechanics, Beijing Institute of Technology, Beijing 100081
Department of Mechanical and Electrical Engineering, Xiamen University, Xiamen 361005

(Received 21 November 2007)

Based on a splitting method and a composition method, we construct some structure-preserving algorithms with first-order, second-order and fourth-order accuracy for a Lorenz system. By using the Liouville's formula, it is proven that the structure-preserving algorithms exactly preserve the volume of infinitesimal cube for the Lorenz system. Numerical experimental results illustrate that for the conservative Lorenz system, the qualitative behaviour of the trajectories described by the classical explicit fourth-order Runge–Kutta (RK4) method and the fifth-order Runge–Kutta–Fehlberg (RKF45) method is wrong, while the qualitative behaviour derived from the structure-preserving algorithms with different orders of accuracy is correct. Moreover, for the small dissipative Lorenz system, the norm errors of the structure-preserving algorithms in phase space are less than those of the Runge–Kutta methods.

PACS: 05.45.Pq, 02.60.Lj

In 1963, in the study of thermal convection motion of atmosphere, Lorenz[1] obtained a three-dimensional system of nonlinear ordinary differential equations, which is known as the Lorenz system (or Lorenz equations). For convenience of discussions, we multiply a parameter $\varepsilon \geq 0$ to each dissipative term in the original Lorenz equations and obtain

$$\begin{aligned} \dot{x} &= \sigma(-\varepsilon x + y), \\ \dot{y} &= \gamma x - xz - \varepsilon y, \\ \dot{z} &= xy - \varepsilon b z, \end{aligned} \tag{1}$$

where σ, γ and b are positive constants. The system expressed by Eqs. (1) here is still called the Lorenz system.

When $\sigma = 10$, $\gamma = 28$, $b = 8/3$ and $\varepsilon = 1$, these seemly simple equations show very complex dynamic behaviour, namely chaotic phenomena. So far, the classical explicit fourth-order Runge–Kutta (RK4) method and the fifth-order Runge–Kutta-Fehlberg (RKF45) method with error controls have been frequently used to solve nonlinear ordinary differential equations. The RK4 and the RKF45 methods are easier for computations and more precise solutions can be obtained in a short time.

Since the appearance of symplectic algorithms,[2] structure-preserving algorithms for ordinary differential equations have attracted much attention. In Ref. [3], Hairer *et al.* described about structure-preserving algorithms in detail. Undoubtedly, the Runge–Kutta methods play an important role in the development of numerical algorithms for ordinary differential equations because these algorithms skillfully avoid the difficulties of Taylor series methods for higher-order derivatives and achieve the same accuracy. However, the explicit Runge–Kutta methods essentially are local approaches and cannot describe the long-time qualitative and global properties of the system.

Completely different from the construction idea of the Runge–Kutta methods, a splitting method was proposed later in 1959 by Trotter.[4] Splitting methods are closely connected with modern mathematical theories of ordinary differential equations. The concepts of flow and one-parameter group of transformations are reflected. Consequently, theories of Lie group and Lie algebra can be used to perform error analysis for results from splitting methods.

Closely connecting with splitting methods is composition methods. Concerning the latter, more representative works are those of Suzuki,[5] Yoshida,[6] Qin and Zhu.[7] Especially, in 1995, McLachlan[8] established the connection of composition methods and splitting methods, and provided the theory of how to construct higher-order methods.

In this Letter, a splitting method and a composition method are adopted to construct structure-preserving algorithms for the Lorenz system and the algorithms are compared with Runge–Kutta methods.

A vector field $\boldsymbol{f}$ can be expressed as the sum of a gradient vector field $\boldsymbol{f}_G : \boldsymbol{R}^n \to \boldsymbol{R}^n$ and a source-free vector field $\boldsymbol{f}_H : \boldsymbol{R}^n \to \boldsymbol{R}^n$,[9] i.e. $\boldsymbol{f} = \boldsymbol{f}_G + \boldsymbol{f}_H$. Since $\boldsymbol{f}_G$ is a gradient field, it can be expressed as a gradient of a scalar function, i.e. $\boldsymbol{f}_G = -I\nabla G(\boldsymbol{u})$, where I is a unit matrix and $G \in C^\infty(\boldsymbol{R}^n, \boldsymbol{R})$. $\boldsymbol{f}_H$ can be decomposed into the sum of Hamiltonian vector fields, namely $\boldsymbol{f}_H = \sum J_i \nabla H_i$ $(1 \leq i \leq n)$, where J_i is a skew-symmetric matrix and H_i is the Hamilto-

*Supported by the National Natural Science Foundation of China under Grant No 10572021, and the Doctoral Programme Foundation of Institute of Higher Education of China under Grant 20040007022.

** To whom correspondence should be addressed. Email: gangtq@126.com

© 2008 Chinese Physical Society and IOP Publishing Ltd

nian function. In a two-dimensional case, $\boldsymbol{f}_H$ is simply a Hamiltonian vector field.

For the Lorenz system, let $\boldsymbol{u} = [x, y, z]^T$, then Eqs. (1) can be decomposed into three parts:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} -\varepsilon\sigma x \\ -\varepsilon y \\ -\varepsilon b z \end{bmatrix} = \begin{bmatrix} -\partial G/\partial x \\ -\partial G/\partial y \\ -\partial G/\partial z \end{bmatrix} = \boldsymbol{f}^{[1]}(\boldsymbol{u}), \tag{2}$$

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} 0 \\ -xz \\ xy \end{bmatrix} = \begin{bmatrix} 0 \\ -\partial H_1/\partial z \\ -\partial H_1/\partial y \end{bmatrix} = \boldsymbol{f}^{[2]}(\boldsymbol{u}), \tag{3}$$

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} \sigma y \\ \gamma x \\ 0 \end{bmatrix} = \begin{bmatrix} \partial H_2/\partial x \\ -\partial H_2/\partial y \\ 0 \end{bmatrix} = \boldsymbol{f}^{[3]}(\boldsymbol{u}). \tag{4}$$

where $G = \varepsilon(\sigma x^2 + y^2 + bz^2)/2$, $H_1 = -x(y^2 + z^2)/2$, $H_2 = (-\gamma x^2 + \sigma y^2)/2$. Thus Eqs. (1) can be rewritten as

$$\dot{\boldsymbol{u}} = \boldsymbol{f}(\boldsymbol{u}) = \boldsymbol{f}^{[1]}(\boldsymbol{u}) + \boldsymbol{f}^{[2]}(\boldsymbol{u}) + \boldsymbol{f}^{[3]}(\boldsymbol{u}). \tag{5}$$

When $\varepsilon = 0$, the Lorenz system becomes a source-free system, which is called a conservative Lorenz system, i.e.

$$\dot{\boldsymbol{u}} = \boldsymbol{f}^{[2]}(\boldsymbol{u}) + \boldsymbol{f}^{[3]}(\boldsymbol{u}). \tag{6}$$

From Eqs. (6), it is clear that this source-free system is composed of two Hamiltonian systems. Some researchers' studies on Lorenz systems are simply based on Eqs. (6) and the perturbation analysis are applied.[10−12]

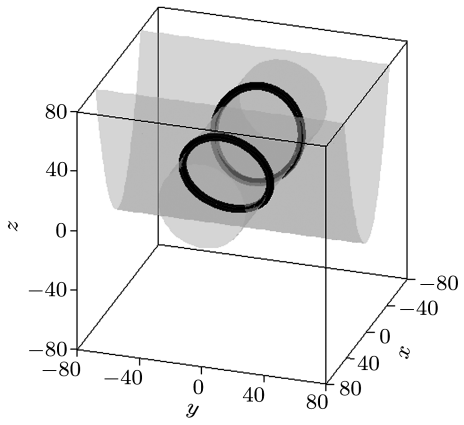


Fig. 1. $|K_1/\sigma| + \gamma > \sqrt{2K_2 + \gamma^2}$.

By derivation, we can obtain three first integrals of Eqs. (6), i.e. $K_1 = x^2/2 - z$, $2K_2 + \gamma^2 = y^2 + (z - \gamma)^2$ and $K_3 = ((\gamma - z)x^2 - y^2)/2$, in which only K_1 and K_2 are large-scale first integrals. Among these three expressions of first integrals, the expression containing K_1 denotes a paraboloid and the expression containing K_2 denotes a cylinder. The intersecting lines of the two surfaces are simply the phase trajectories of Eqs. (6). According to the relative positions of two different surfaces, the trajectories include three cases:

(a) For $|K_1/\sigma| + \gamma > \sqrt{2K_2 + \gamma^2}$, the system has two isolated closed trajectories, as shown in Fig. 1.

(b) For $|K_1/\sigma| + \gamma = \sqrt{2K_2 + \gamma^2}$, the system has a pair of homoclinic trajectories, as shown in Fig. 2.

(c) For $|K_1/\sigma| + \gamma < \sqrt{2K_2 + \gamma^2}$, the system has one closed trajectory, as shown in Fig. 3.

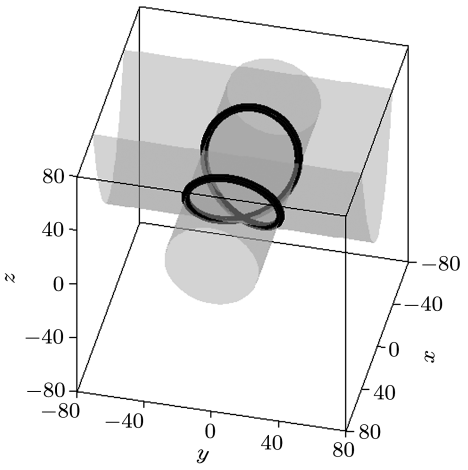


Fig. 2. $|K_1/\sigma| + \gamma = \sqrt{2K_2 + \gamma^2}$.

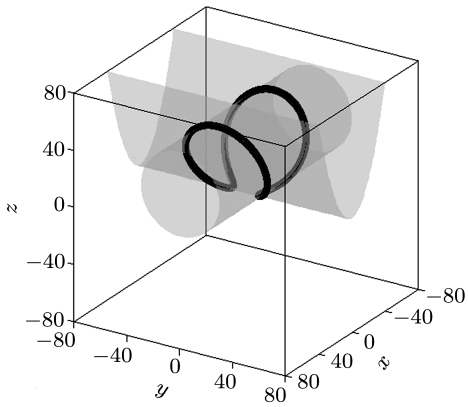


Fig. 3. $|K_1/\sigma| + \gamma < \sqrt{2K_2 + \gamma^2}$.

For the conservative Lorenz system, first of all, we use a splitting method and a composition method to construct the structure-preserving algorithms of Eqs. (6). Since the subsystems (3) and (4) are linear Hamiltonian systems, there correspondingly exist exact flow mappings $\boldsymbol{\varphi}_t^{[2]}$ and $\boldsymbol{\varphi}_t^{[3]}$, respectively, i.e.

$$\boldsymbol{\varphi}_t^{[2]}(\boldsymbol{u}) = \begin{bmatrix} x \\ y\cos(|x|t) - \operatorname{sign}(x) z \sin(|x|t) \\ -\operatorname{sign}(x) y \sin(|x|t) + z\cos(|x|t) \end{bmatrix},$$

$$\boldsymbol{\varphi}_t^{[3]}(\boldsymbol{u}) = \begin{bmatrix} ((x + \mu y)e^{\eta t} + (x - \mu y)e^{-\eta t})/2 \\ ((x/\mu + y)e^{\eta t} + (-x/\mu + y)e^{-\eta t})/2 \\ z \end{bmatrix},$$

where $\mu = \sqrt{\sigma/\gamma}$, $\eta = \sqrt{\sigma\gamma}$,

$$\text{sign}(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0 \end{cases}.$$

It is easy to prove that the determinant values of the Jacobian matrices of $\boldsymbol{\varphi}_t{}^{[2]}$ and $\boldsymbol{\varphi}_t^{[3]}$ are $\det(D_{\boldsymbol{u}}\boldsymbol{\varphi}_t^{[2]}(\boldsymbol{u})) = \det(D_{\boldsymbol{u}}\boldsymbol{\varphi}_t^{[3]}(\boldsymbol{u})) = 1$. Therefore, $\boldsymbol{\varphi}_t^{[2]}$ and $\boldsymbol{\varphi}_t^{[3]}$ preserve the volume of phase flow invariable and their composition also have the same quality, i.e. $\det(D_{\boldsymbol{u}}(\boldsymbol{\varphi}_t^{[2]} \circ \boldsymbol{\varphi}_t^{[3]}(\boldsymbol{u}))) = 1$. Preserving the volume of phase flow invariable is an essential quality of a source-free system. In terms of $\boldsymbol{\varphi}_t^{[2]}$ and $\boldsymbol{\varphi}_t^{[3]}$, the structure-preserving algorithm for Eqs. (6) with first-order accuracy is $\boldsymbol{\psi}_h^1 = \boldsymbol{\varphi}_h^{[2]} \circ \boldsymbol{\varphi}_h^{[3]}$ (where h is the step size) and its adjoint method is $(\boldsymbol{\psi}_h^1)^* = \boldsymbol{\varphi}_h^{[3]} \circ \boldsymbol{\varphi}_h^{[2]}$. Composing the first-order method and its adjoint method, we can obtain the second-order method, i.e. $\boldsymbol{\psi}_h^2 = \boldsymbol{\psi}_{h/2}^1 \circ (\boldsymbol{\psi}_{h/2}^1)^*$. Since $\boldsymbol{\psi}_h^2 = (\boldsymbol{\psi}_h^2)^*$, $\boldsymbol{\psi}_h^2$ is a symmetric approach. Based on $\boldsymbol{\psi}_h^2$, by using the composition method,[3] a higher-order method can be constructed. Taking a fourth-order method as an example, the method is $\boldsymbol{\psi}_h^4 = \boldsymbol{\psi}_{r_1 h}^2 \circ \boldsymbol{\psi}_{r_2 h}^2 \circ \boldsymbol{\psi}_{r_3 h}^2$, in which $r_1 = r_3 = 1/(2 - \sqrt[3]{2})$ and $r_2 = -\sqrt[3]{2}/(2 - \sqrt[3]{2})$. This fourth-order method $\boldsymbol{\psi}_h^4$ is still a structure-preserving algorithm.

For Eqs. (6), numerical computations are performed by using the RK4, the RKF45 and the structure-preserving algorithms with first-order, second-order and fourth-order accuracy obtained above. In these methods, the step size is $h = 0.01$ (for the RKF45, it is only an initial step size) and the simulating time is $t \in [0, 500]$. In addition, in the RKF45, the error control parameter is $\delta = 10^{-5}$. Choose the initial condition values as $x(0) = 0.1$ and $y(0) = z(0) = 0$, which makes values of $|K_1/\sigma| + \gamma$ and $\sqrt{2K_2 + \gamma^2}$ closer. Because $(|K_1/\sigma| + \gamma) - \sqrt{2K_2 + \gamma^2} = 0.0005 > 0$, in this case, according to the analysis of the qualitative behaviour of the phase trajectories for Eqs. (6), the phase trajectories should be two isolated closed ones. For convenience of observation, the phase trajectories derived from different methods are projected on the $x - y$ plane, as shown in Figs. 4–6.

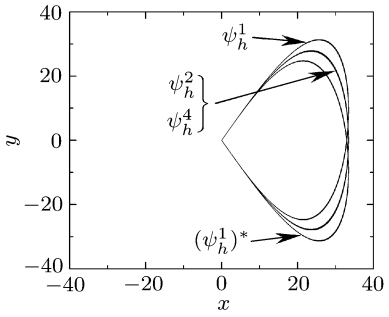


Fig. 4. Trajectories of the structure-preserving algorithms.

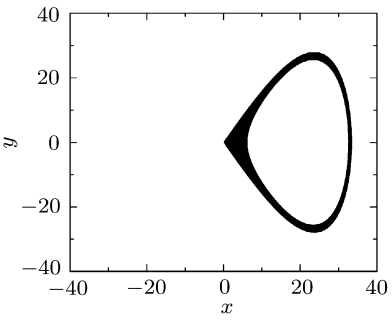


Fig. 5. Trajectory of the RK4.

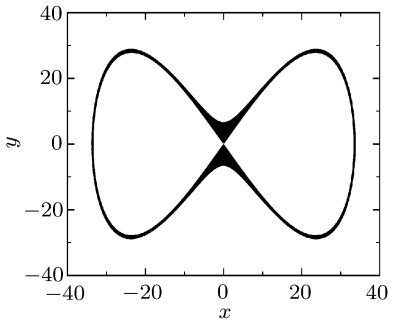


Fig. 6. Trajectory of the RKF45.

Figure 4 illustrates that the qualitative behaviour of the structure-preserving algorithms with different order accuracies is correct and no dissipation exists. Under the precision shown in Fig. 4, the phase trajectories of $\boldsymbol{\psi}_h^2$ and $\boldsymbol{\psi}_h^4$ are overlapped. $\boldsymbol{\psi}_h^1$ and $(\boldsymbol{\psi}_h^1)^*$ are not symmetric algorithms, hence the corresponding trajectories are not symmetric about y-axis. Although the results by the RK4 in Fig. 5 is maintaining within one of the closed trajectories, the dissipation is serious. The trajectory of the RK4 tends to run spirally within the closed trajectory.

In Fig. 6, the trajectory obtained by the RKF45 is developed spirally outward and run from the right closed branch into the left one, which corresponds to $|K_1/\sigma| + \gamma < \sqrt{2K_2 + \gamma^2}$. Therefore, from the point of qualitative analysis, this is a more serious mistake.

When $\varepsilon > 0$, Eqs. (5) denote a dissipative Lorenz system and correspondingly we can construct the structure-preserving algorithms in accordance with the following steps. The exact flow mapping of Eqs. (2) is $\boldsymbol{\varphi}_t^{[1]}(\boldsymbol{u}) = [e^{-\varepsilon\sigma t}x,\ e^{-\varepsilon t}y,\ e^{-\varepsilon b t}z]^T$. The first-order method for Eqs. (5) is $\boldsymbol{\phi}_h^1 = \boldsymbol{\varphi}_h^{[1]} \circ \boldsymbol{\varphi}_h^{[2]} \circ \boldsymbol{\varphi}_h^{[3]}$ and the adjoint method of $\boldsymbol{\phi}_h^1$ is $(\boldsymbol{\phi}_h^1)^* = \boldsymbol{\varphi}_h^{[3]} \circ \boldsymbol{\varphi}_h^{[2]} \circ \boldsymbol{\varphi}_h^{[1]}$. Considering higher-order methods for the dissipative Lorenz system, the construction method is similar to that of $\boldsymbol{\psi}_h^4$. Thus, we obtain the second-order method $\boldsymbol{\phi}_h^2 = \boldsymbol{\phi}_{h/2}^1 \circ (\boldsymbol{\phi}_{h/2}^1)^*$ and the fourth-order method $\boldsymbol{\phi}_h^4 = \boldsymbol{\phi}_{r_1 h}^2 \circ \boldsymbol{\phi}_{r_2 h}^2 \circ \boldsymbol{\phi}_{r_3 h}^2$ of the Lorenz system. In method $\boldsymbol{\phi}_h^4$, r_1, r_2 and r_3 are the same as those in $\boldsymbol{\psi}_h^4$. For the conservative Lorenz system, $\boldsymbol{\psi}_h^1$, $\boldsymbol{\psi}_h^2$ and $\boldsymbol{\psi}_h^4$ can be considered as $\boldsymbol{\phi}_h^1$, $\boldsymbol{\phi}_h^2$ and $\boldsymbol{\phi}_h^4$ respectively for the special case $\varepsilon = 0$.

Given an initial value $\boldsymbol{u}_0$, we use $\boldsymbol{\varphi}_t = \boldsymbol{\varphi}(t; \boldsymbol{u}_0)$ to denote the exact flow mapping of $\dot{\boldsymbol{u}} = \boldsymbol{f}(\boldsymbol{u})$ (Eqs. 5) in $\boldsymbol{R}^3$. According to Liouville's formula,[13] we have

$$\det(D_{\boldsymbol{u}}\boldsymbol{\varphi}_t) = \exp\Big(\int_0^t (\nabla \cdot \boldsymbol{f}(\boldsymbol{u}))ds\Big)$$

$$= \exp(-\varepsilon(\sigma + b + 1)t),$$

where $\det(D_{\boldsymbol{u}}\boldsymbol{\varphi}_t)$ is the determinant of the Jacobian matrix of $\boldsymbol{\varphi}_t$ and $\nabla \cdot \boldsymbol{f}(\boldsymbol{u})$ is the divergence of $\boldsymbol{f}(\boldsymbol{u})$. In Ref. [13], $\det(D_{\boldsymbol{u}}\boldsymbol{\varphi}_t)$ is called the volume of an infinitesimal cube. Let $\boldsymbol{\phi}_h^{SPA}$ represent all the structure-preserving algorithms derived above for $\dot{\boldsymbol{u}} = \boldsymbol{f}(\boldsymbol{u})$. If h is replaced by time t, $\boldsymbol{\phi}_t^{SPA}$ is also a one-parameter group of transformations. In terms of the backward error analysis theory of numerical methods for ordinary differential equations,[3] $\boldsymbol{\phi}_t^{\rm SPA}$ is an exact flow mapping of the modified equations $\dot{\boldsymbol{u}} = \tilde{\boldsymbol{f}}_{\rm SPA}(\boldsymbol{u}, h)$ of Eqs. (5). Then, $\det(D_{\boldsymbol{u}}\boldsymbol{\phi}_t^{\rm SPA}) = \det(D_{\boldsymbol{u}}\boldsymbol{\varphi}_t)$ can be proved. Hence, in phase space, the structure-preserving algorithms exactly preserve the volume of infinitesimal cube for the Lorenz system. In a similar way, let $\boldsymbol{\phi}_t^{\rm RK}$ represent exact flows of the RK4 and the RKF45, it is easy to prove that $\det(D_{\boldsymbol{u}}\boldsymbol{\phi}_t^{RK}) \neq \det(D_{\boldsymbol{u}}\boldsymbol{\varphi}_t)$. Therefore, for simulations of long-time qualitative behaviour of the Lorenz system, the RK4 and the RKF45 methods are not suitable.

When the trajectory of the Lorenz system shows a chaotic phenomenon, the solution is very sensitive to the order and the step size of the numerical methods. In this case, there is no sense to discuss the solutions accuracy for a given initial condition. Thus, in the following, we can simply perform numerical experiments under the condition $0 < \varepsilon \ll 1$.

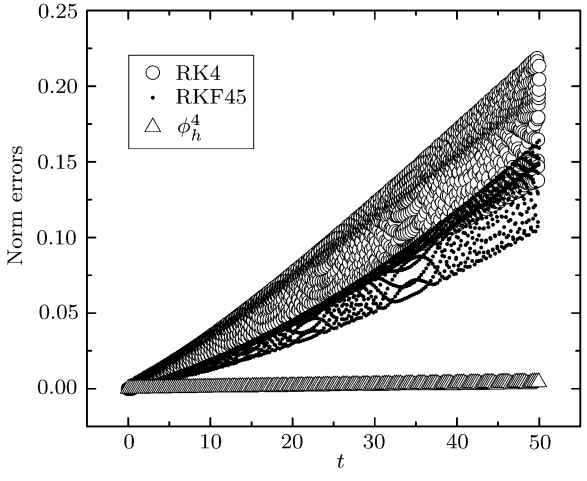


Fig. 7. Norm errors of solutions by different methods.

For $0 < \varepsilon \ll 1$, the system expressed by Eqs. (5) is called a small dissipative one, with $\nabla \cdot \boldsymbol{f}(\boldsymbol{u}) = -\varepsilon(\sigma + b + 1) \ll 1$. For numerical experiments, we use $\varepsilon = 0.001$, $\sigma = 10$, $\gamma = 28$, $b = 8/3$, $h = 0.006$ and $t \in [0, 50]$. Since the computation time is relatively shorter, we can use results of the dverk78 method in Maple (a computer algebra language) as the exact solutions. The dverk78 can find a numerical solution using a seventh-eighth order continuous Runge–Kutta method. It is capable of working in arbitrary precision based on the setting of Maple and can be used to obtain high accuracy solutions. However, it takes more time than other methods. The corresponding error is defined as $\|\boldsymbol{u}_{\rm dverk78} - \boldsymbol{u}_n\|/\|\boldsymbol{u}_{\rm dverk78}\|$, which denotes the norm error of solutions in the phase space. Figure 7 shows the norm errors of solutions by $\boldsymbol{\phi}_h^4$, the RK4 and the RKF45. It can be concluded that the norm errors by $\boldsymbol{\phi}_h^4$ are much smaller than those by the RK4 and the RKF45, i.e. the results of the structure-preserving algorithm $\boldsymbol{\phi}_h^4$ are better than those of the RK4 and RKF45.

From all the above discussions, it is clear that the structure-preserving algorithms constructed in this study exactly preserve the volume of infinitesimal cube for the Lorenz system. For both the conservative and the dissipative Lorenz system, the structure-preserving algorithms can correctly reflect the long-time evolution of the qualitative behaviour. Contrarily, the computation results of the explicit Runge–Kutta methods show wrong qualitative behaviour and they are not suitable for simulations of long-time qualitative behaviours of the Lorenz system. Furthermore, for the small dissipative Lorenz system, we perform norm error analysis of different methods. It reaches the conclusion that the norm errors of the structure-preserving algorithm are evidently smaller than those of the explicit Runge–Kutta methods.

References

[1] Lorenz E N 1963 *J. Atmospheric Sci.* **20** 130
[2] Feng K 1984 *Proceedings of International Symposium on Differential Geometry and Differential Equations* (Beijing: Science Press)
[3] Hairer E, Lubich C and Wanner G 2002 *Geometric Numerical Integration* (Berlin: Springer)
[4] Trotter H F 1959 *Proc. Am. Math. Soc.* **10** 545
[5] Suzuki M 1990 *Phys. Lett.* A **146** 319
[6] Yoshida H 1990 *Phys. Lett.* A **150** 262
[7] Qin M Z and Zhu W J 1992 *Computing.* **47** 309
[8] McLachlan R I 1995 *SIAM J. Sci. Comput.* **16** 151
[9] Mendes R V and Taborda D 1981 *J. Math. Phys.* **22** 1420
[10] Robbins K 1979 *SIAM J. Appl. Math.* **36** 457
[11] Sparrow C 1982 *The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors* (New York: Springer)
[12] Jackson E A 1990 *Perspectives of Nonlinear Dynamics* (Cambridge: Cambridge University Press)
[13] Robinson C R 2007 *An Introduction to Dynamical Systems: Continuous and Discrete* (Beijing: China Machine Press) (in Chinese)