Computers Math. Applic. Vol. 26, No. 4, pp. 33–42, 1993
Printed in Great Britain.

0898-1221/93 \$6.00 + 0.00

Volume-Preserving Schemes and Numerical Experiments

M.-Z. QIN AND W.-J. ZHU
State Key Laboratory of Scientific and Engineering Computing
Computing Center, Academia Sinica, Beijing 100080

(Received and accepted November 1992)

Abstract—In this paper, we discuss the conditions for Euler midpoint rule to be volume-preserving and present explicit volume preserving schemes. Some numerical experiments are done to test these schemes.

1. INTRODUCTION

It's well-known that classical difference schemes such as Euler midpoint rule and fourth-order Runge-Kutta method can be applied to any kind of dynamical systems neglecting the structure of the systems, but the numerical results may be sometimes unsatisfactory. Recently, Feng Kang pointed out that for dynamical systems with specific geometric structure, it is natural and mandatory to require the difference scheme to be structure-preserving, [1]. That is to say for source-free systems, we should use volume-preserving schemes. In this paper, we point out that Euler midpoint rule is only volume-preserving under some strict conditions, and we will present explicit volume-preserving scheme which is also structure-preserving for any dynamical systems.

2. CONDITIONS FOR EULER METHOD TO BE VOLUME-PRESERVING

Let $x = (x_1, x_2, \dots, x_N)' \in R^N$ and $f(x) = (f_1(x), f_2(x), \dots, f_N(x)), f_i(x) : R^N \to R^N$, then the dynamical system

$$\frac{dx}{dt} = f(x) \tag{1}$$

is source-free (i.e., divergence-free) when $\sum_{i=1}^{N} \frac{\partial f_i}{\partial x_i} = 0$ (i.e., $\operatorname{div} f(x) = 0$), [2–4]. The flow of a source-free system is volume-preserving, i.e., $\det(e_f^t(x))_* = 1, \forall x, t$, where e_f^t denotes the flow of (1) and $(e_f^t(x))_*$ the Jacobian matrix of e_f^t at x. So volume-preserving schemes are required for computing the numerical solution of (1). That means the schemes should satisfy $\det\left(\frac{\partial x_{n+1}}{\partial x_n}\right)=1$, if y_n denotes the numerical solution at step n.

Let's consider the Euler midpoint rule [5]

$$x_{n+1} = x_n + \tau f\left(\frac{x_{n+1} + x_n}{2}\right), \tag{2}$$

Supported by National Natural Science Foundation of China, finished August 1992.

Typeset by $\mathcal{A}_\mathcal{M}\mathcal{S}$-TEX where τ is the step-size in t. We then have

$$\begin{aligned}
&\frac{\partial x_{n+1}}{\partial x_n} = I_N + \tau Df\left(\frac{x_{n+1}+x_n}{2}\right)\left(\frac{1}{2}\frac{\partial x_{n+1}}{\partial x_n} + \frac{1}{2}I_N\right) \\
\Longrightarrow &\frac{\partial x_{n+1}}{\partial x_n} = \frac{I_N + \frac{\tau}{2}Df(x^*)}{I_N - \frac{\tau}{2}Df(x^*)},
\end{aligned} \tag{3}$$

where $Df = f_x = \dfrac{\partial f}{\partial x} = B = \{b_{ij}\}, x^* = \dfrac{x_{n+1}+x_n}{2}$.

The condition $\det\left(\dfrac{\partial x_{n+1}}{\partial x_n}\right) = 1$ now requires $\dfrac{|I_N + \frac{\tau}{2}Df(x^*)|}{|I_N - \frac{\tau}{2}Df(x^*)|} = 1$. Let

$$P(\lambda) = |Df(x^*) - \lambda I_N|, \tag{4}$$

be the characteristic matrix of $Df(x^*)$, since

$$\frac{|I_N + \frac{\tau}{2}Df(x^*)|}{|I_N - \frac{\tau}{2}Df(x^*)|} = \frac{|\frac{\tau}{2}(Df(x^*) + \frac{2}{\tau}I_N)|}{|-\frac{\tau}{2}(Df(x^*) - \frac{2}{\tau}I_N)|} = \frac{(\frac{\tau}{2})^N|Df(x^*) + \frac{2}{\tau}I_N|}{(-\frac{\tau}{2})^N|Df(x^*) - \frac{2}{\tau}I_N|} = (-1)^N\frac{P(\frac{2}{\tau})}{P(-\frac{2}{\tau})},$$

we then get the condition for scheme (2) to be volume-preserving

$$P(\lambda) = (-1)^N P(-\lambda). \tag{5}$$

Let's consider some particular cases of N to show that scheme (2) is not always volume-preserving.

CASE 1. $N = 2$. In this case, we have

$$P(\lambda) = \lambda^2 - (b_{11} + b_{22})\lambda + b_{12}b_{21}.$$

Since $\displaystyle\sum_{i=1}^{N}\frac{\partial f_i}{\partial x_i} = 0$, i.e., $\operatorname{tr}(B) = 0$, then $P(\lambda) = \lambda^2 + b_{12}b_{21}$ and $P(-\lambda) = P(\lambda)$, thus the scheme (2) is volume-preserving for source-free systems of dimension 2.

CASE 2. $N = 3$. We now have

$$\begin{aligned}
P(\lambda) &= -\lambda^3 + (b_{11} + b_{22} + b_{33})\lambda^2 - c\lambda + |B| \\
&= -\lambda^3 - c\lambda + |B|,
\end{aligned}$$

where

$$c = \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} + \begin{vmatrix} b_{22} & b_{23} \\ b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{13} \\ b_{31} & b_{33} \end{vmatrix}.$$

The volume-preserving condition is now $|B| = 0$.

For example, when system (1) takes the form

$$\begin{cases} \dfrac{dx}{dt} = cy - bz \\ \dfrac{dy}{dt} = az - cx\,, & a, b, c \in R, \\ \dfrac{dz}{dt} = bx - ay \end{cases}$$

we have $|B| = 0$. For this dynamical system, centered Euler method is volume-preserving.

To study the general cases, we need the following well-known lemma.

LEMMA 1.1. *Let $P(\lambda)$ be the characteristic polynomial of matrix $A_{N\times N}$, then*

$$P(\lambda) = |A - \lambda I_N| = (-1)^N(\lambda^N - P_1\lambda^{N-1} + P_2\lambda^{N-2} + \cdots + (-1)^N P_N),$$

where

$$\begin{aligned} P_1 &= \sum_i^N a_{ii} = \mathrm{tr}(A), \\ P_2 &= \sum_{i<j}^N \begin{vmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{vmatrix}, \\ P_3 &= \sum_{i<j<k}^N \begin{vmatrix} a_{ii} & a_{ij} & a_{ik} \\ a_{ji} & a_{jj} & a_{jk} \\ a_{ki} & a_{kj} & a_{kk} \end{vmatrix}, \\ &\cdots\cdots \\ P_N &= |A|. \end{aligned} \tag{6}$$

Using Lemma 1.1, we can discuss the case when $N = 4$.

CASE 3. $N = 4$. In this case,

$$P(\lambda) = \lambda^4 - P_1\lambda^3 + P_2\lambda^2 - P_3\lambda + |B|.$$

Since $P_1 = \mathrm{tr}(B) = 0$, then $P(-\lambda) = (-1)^4 P(\lambda)$ requires $P_3 = 0$.

We note, when N increases, more and more conditions are required for scheme (2) to be volume-preserving, and it seems impossible to satisfy all these conditions. But fortunately, for the special case when system (1) is Hamiltonian, i.e.,

$$f = J\nabla H, \quad J = \begin{pmatrix} 0 & -I_K \\ I_K & 0 \end{pmatrix}, \qquad N = 2K.$$

Scheme (2) is volume-preserving. This is because the Hamiltonian system is source-free and Df is an infinitesimal symplectic matrix. For infinitesimal symplectic matrix, we have the following lemma (see [6]).

LEMMA 1.2. *If M is an infinitesimal symplectic matrix and λ is an egienvalue of M, so are $-\lambda, \bar{\lambda}, -\bar{\lambda}$.*

From Lemma 1.2, we know $P(-\lambda) = (-1)^{2K}P(\lambda)$ is valid when system (1) is Hamiltonian, so Euler method is volume-preserving for Hamiltonian systems. In fact, the method is even symplectic for Hamiltonian systems, that is to say it also preserves the symplectic structure of Hamiltonian systems, which is a much stronger property than volume-preserving.

3. SEPARABLE SYSTEMS AND VOLUME-PRESERVING EXPLICIT METHODS

In this section, we consider a special kind of source-free systems called separable systems. System (1) is separable iff

$$\frac{dx_i}{dt} = f(x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_N), \qquad i = 1, 2, \ldots, N. \tag{7}$$

We can divide the system (7) into N source-free systems. The first order explicit Euler method can be applied to them to get the exact solutions of them, i.e., the phase flows of them. Using the composition method, we can construct first order explicit volume-preserving scheme for system (7). The adjoint of this scheme is got from the implict Euler method and is also explicit. Composing these two schemes, we get a reversible explict volume-prserving scheme of order 2. This process can be expressed by formal power series as shown in the following.

From [1] we know the flow of (1) can be represented by power series

$$e_f^\tau = 1_N + \sum_{k=1}^{\infty} \tau^k e_{k,f}, \quad e_{k,f} : R^N \longrightarrow R^N, \qquad e_{k,f} = \frac{1}{k!} f^{*k} 1_N,$$

where f^* denotes the first order differential operator $f^* = \sum_{i=1}^{N} f_i \frac{\partial}{\partial x_i}$ and $f^{*2} = f^* f^*, f^{*3} = f^* f^* f^*, \ldots$ 1_N is the identity vector function, $1_N(x) = x$. The concatenation of two flows e_A^τ and e_B^τ, where A, B are vectors like f, is a formal power series corresponding to a formal vector field C^τ, which is given by the following formula. For simplicity, we just write out first several terms.

$$\begin{aligned} &e_A^\tau \circ e_B^\tau = e_{C^\tau}^\tau, \\ &C^\tau = A + B + \frac{\tau}{2}[A, B] + o(\tau^2), \end{aligned} \tag{8}$$

where $[A, B] = A_* B - B_* A$ is the Lie bracket of A and B, A_* denotes the Jacobian matrix of A. For details about formal power series and formal vector fields, refer to [1].

We now separate (1) into N integrable systems

$$\frac{dx}{dt} = a_i(x), \quad a_i = (0, \vdots, 0, f_i, 0, \ldots, 0)^\top, \qquad i = 1, 2, \ldots, N. \tag{9}$$

These integrable systems have flows

$$e_{a_i}^\tau = 1_N + \sum_{k=1}^{\infty} \tau^k e_{k,a_i}, \qquad i = 1, \ldots, N. \tag{10}$$

Since we have $a_i^{*k} 1_N(x) = a_i^{*k} x = 0$ when $k \geq 2$, then

$$e_{a_i}^\tau(x) = x + \sum_{k=1}^{\infty} \tau^k e_{k,a_i}(x) = x + \sum_{k=1}^{\infty} \frac{\tau^k}{k!} a_i^{*k} 1_N(x) = x + \tau a_i(x). \tag{11}$$

Using the formula (8), we can find

$$e_{a_N}^\tau \circ e_{a_{N-1}}^\tau \circ \cdots \circ e_{a_2}^\tau \circ e_{a_1}^\tau = e_{f+o(\tau)}^\tau. \tag{12}$$

From [1] we know this means the concatenation $e_{a_N}^\tau \circ e_{a_{N-1}}^\tau \circ \cdots \circ e_{a_2}^\tau \circ e_{a_1}^\tau$ approximates the flow e_f^τ to the first order of τ. Because the systems of (9) are all source free, their flows are all volume-preserving and the concatenation of them remains volume-preserving since

$$\begin{aligned} &\det\left(\left(e_{a_N}^\tau \circ e_{a_{N-1}}^\tau \circ \cdots \circ e_{a_2}^\tau \circ e_{a_1}^\tau\right)(x)\right)_* \\ =&\det(e_{a_N}^\tau(x^{N-1}))_* \cdot \det(e_{a_{N-1}}^\tau(x^{N-2}))_* \cdot \cdots \cdot \det(e_{a_1}^\tau(x^0))_* \\ =&1, \end{aligned}$$

where $x^0 = x, x^1 = e_{a_1}^\tau(x^0), \ldots, x^{N-1} = e_{a_{n-1}}^\tau(x^{n-2}), x^N = e_{a_N}^\tau(x^{N-1})$.

Thus, we get the volume-preserving scheme $e_{a_N}^\tau \circ e_{a_{N-1}}^\tau \circ \cdots \circ e_{a_2}^\tau \circ e_{a_1}^\tau$, it's an explicit scheme since $e_{a_i}^\tau (i = 1, \ldots, N)$ are flows of integrable systems which can be written as (11). From [7], we know the concatenation of $e_{a_N}^\tau \circ e_{a_{N-1}}^\tau \circ \cdots \circ e_{a_2}^\tau \circ e_{a_1}^\tau$ with its adjoint $e_{a_1}^\tau \circ e_{a_2}^\tau \circ \cdots \circ e_{a_{N-1}}^\tau \circ e_{a_N}^\tau$ produces a reversible scheme

$$e_{a_N}^{\tau/2} \circ e_{a_{N-1}}^{\tau/2} \circ \cdots \circ e_{a_2}^{\tau/2} \circ e_{a_1}^{\tau/2} \circ e_{a_1}^{\tau/2} \circ e_{a_2}^{\tau/2} \circ \cdots \circ e_{a_{N-1}}^{\tau/2} \circ e_{a_N}^{\tau/2}$$

of second order.

We write out the reversible explicit volume-preserving scheme of second order for system (1) with $N = 3$ as an example.

$$\begin{aligned}
x_1^{(1/2)} &= x_1^{(0)} + \frac{\tau}{2} f_1(x_2^{(0)}, x_3^{(0)}) \\
x_2^{(1/2)} &= x_2^{(0)} + \frac{\tau}{2} f_2(x_1^{(1/2)}, x_3^{(0)}) \\
x_3^{(1/2)} &= x_3^{(0)} + \frac{\tau}{2} f_3(x_1^{(1/2)}, x_2^{(1/2)}) \\
x_3^{(1)} &= x_3^{(1/2)} + \frac{\tau}{2} f_3(x_1^{(1/2)}, x_2^{(1/2)}) \\
x_2^{(1)} &= x_2^{(1/2)} + \frac{\tau}{2} f_2(x_1^{(1/2)}, x_3^{(1)}) \\
x_1^{(1)} &= x_1^{(1/2)} + \frac{\tau}{2} f_1(x_2^{(1)}, x_3^{(1)}),
\end{aligned} \tag{13}$$

where $(x_1^{(0)}, x_2^{(0)}, x_3^{(0)})$ denotes the numerical solution at some step and $(x_1^{(1)}, x_2^{(1)}, x_3^{(1)})$ the next.

It is easy to construct volume-preserving schemes for linear systems and we will not discuss this problem here. For separable systems, we have the volume-preserving scheme (13) which is good enough. But for non-separable systems, it seems very difficult to construct praticable volume-preserving schemes. We give out some volume-preserving schemes for some special systems without proofs.

SCHEME 1. For system

$$\begin{cases} \dfrac{dx}{dt} = u(x,y) \\ \dfrac{dy}{dt} = v(x,y) \\ \dfrac{dz}{dt} = w(x,y,z) \end{cases},$$

the scheme

$$\begin{cases} x_{n+1} = x_n + \tau u(x_n, y_n) \\ y_{n+1} = y_n + \tau v(x_n, y_n) \\ z_{n+1} = \dfrac{z_n + \tau w(x_n, y_n, z_{n+1})}{1 + u_x(x_n, y_n) v_y(x_n, y_n) - u_y(x_n, y_n) v_x(x_n, y_n)} \end{cases}$$

is volume-preserving and of first order.

SCHEME 2. For system

$$\begin{cases} \dfrac{dx}{dt} = u(z) \\ \dfrac{dy}{dt} = v(x,y,z) \\ \dfrac{dz}{dt} = w(x,y,z) \end{cases},$$

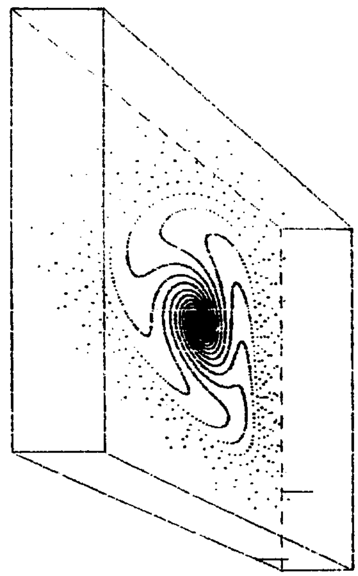
the scheme

$$\begin{cases} x_{n+1} = x_n + \tau u(z_{n+1}) \\ y_{n+1} = y_n + \tau v(x_n, y_n, z_{n+1}) \\ z_{n+1} = z_n - \tau w(x_n, y_n, z_{n+1}) \end{cases}$$

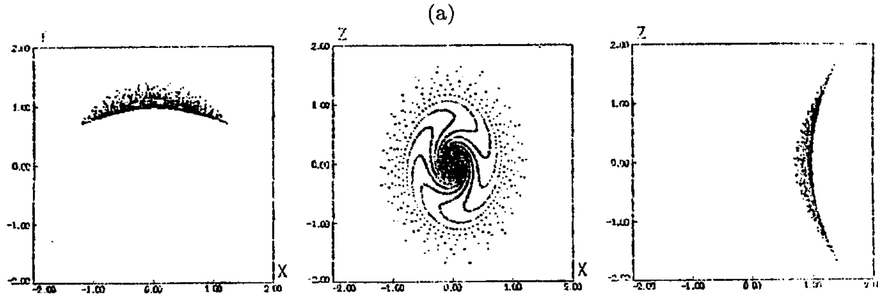
is volume-preserving and of first order.

SCHEME 3. For system

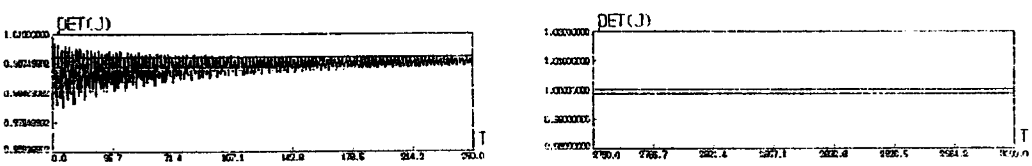
$$\begin{cases} \dfrac{dx}{dt} = u(x,z) \\ \dfrac{dy}{dt} = v(x,z) \\ \dfrac{dz}{dt} = w(x,y,z) \end{cases},$$



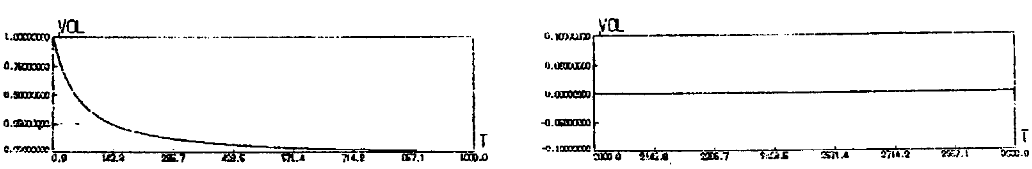
(a)



(b)



(c)



(d)

Figure 1.

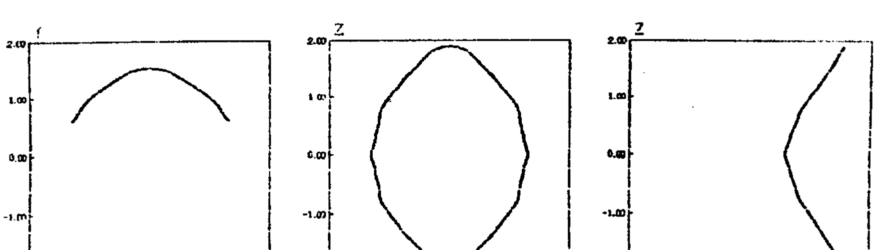
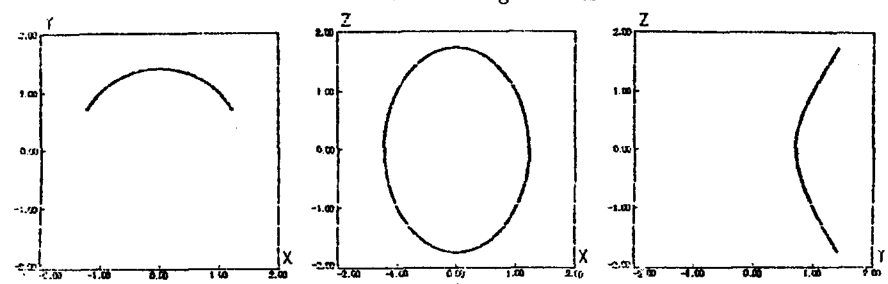
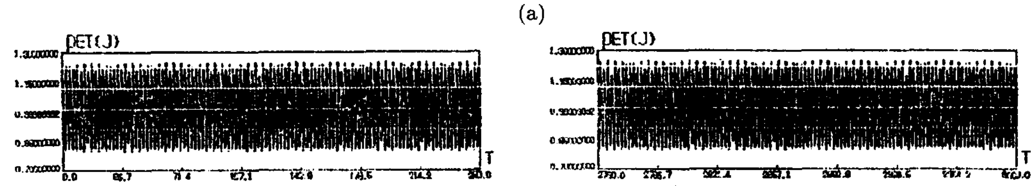


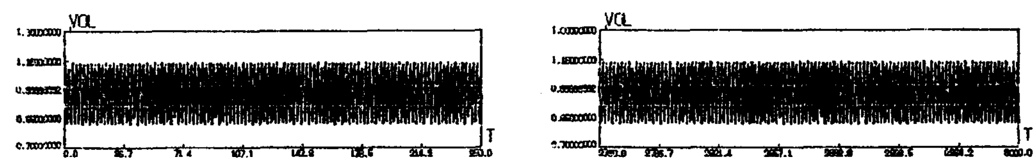
Figure 2.



(a)



(b)



(c)

Figure 3.

the scheme

$$\begin{cases} x_{n+1} = x_n + \tau u(x_n, z_{n+1}) \\ y_{n+1} = y_n + \tau v(x_n, z_{n+1}) \\ z_{n+1} = z_n - \tau w(x_n, y_n, z_{n+1}) \end{cases}$$

is volume-preserving and of first order.

4. NUMERICAL EXPERIMENTS

NUMERICAL EXPERIMENT 1. We first consider the following system

$$\begin{cases} \dot{x} = yz \\ \dot{y} = -xz \\ \dot{z} = -k^2xy. \end{cases} \tag{14}$$

It's a source-free system and has two integral invariants

$$x^2 + y^2 = C_1, \tag{15.1}$$
$$k^2x^2 + z^2 = C_2, \tag{15.2}$$

from which we know

$$k^2y^2 - z^2 = C_3 \tag{15.3}$$

is also invariant under the flow of (14). So the phase trajectory of this system is closed and the projection of it on to the xy plane will be a circle, on to the xz plane an ellipse and on to the yz plane a hyperbola. In our numerical experiments, we always take the initial values as $x_0 = y_0 = z_0 = 1.0$ and $k = \sqrt{2}$, so (15) will be

$$x^2 + y^2 = 2.0 \tag{16.1}$$
$$2x^2 + z^2 = 3.0 \tag{16.2}$$
$$2y^2 - z^2 = 1.0, \tag{16.3}$$

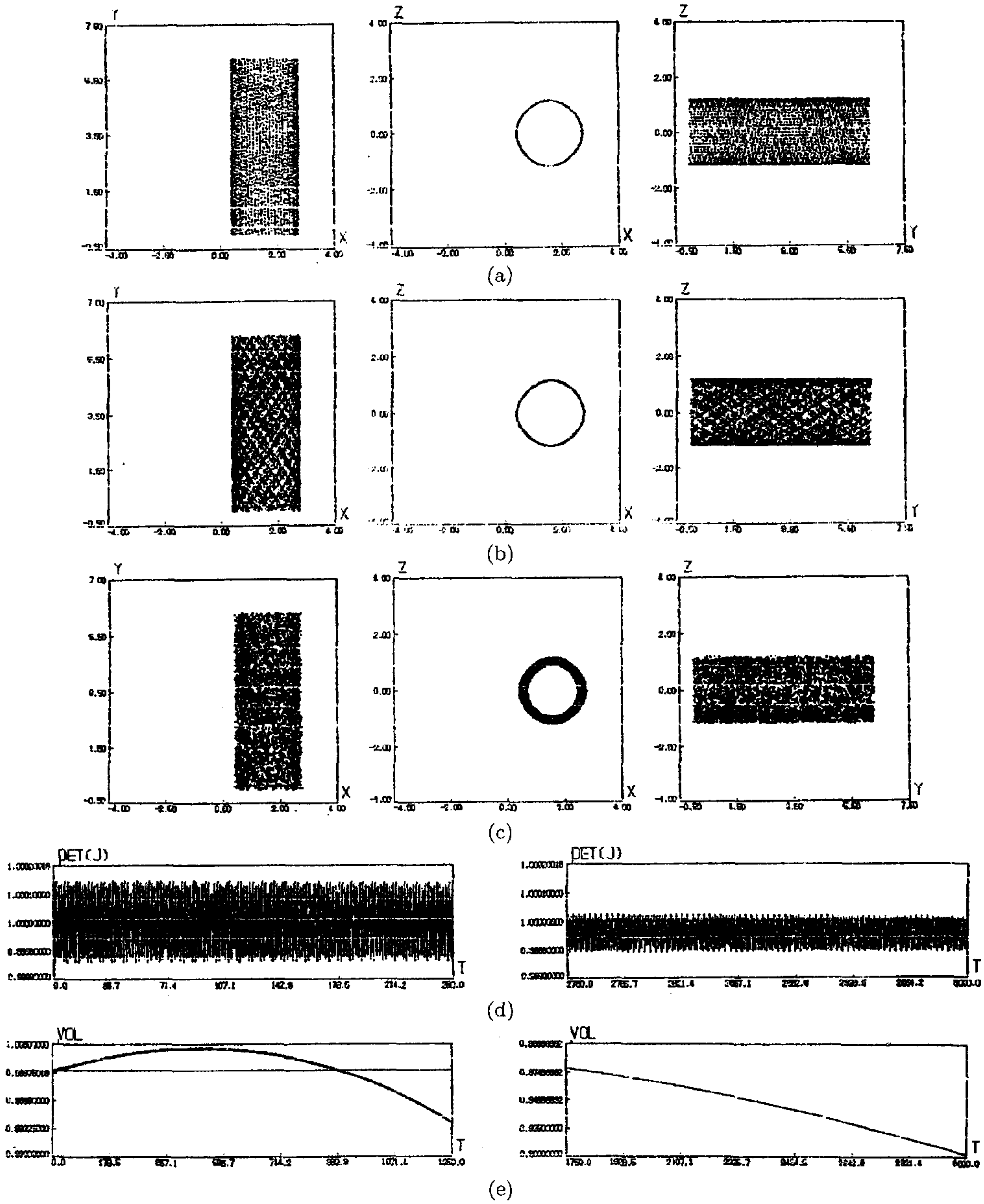
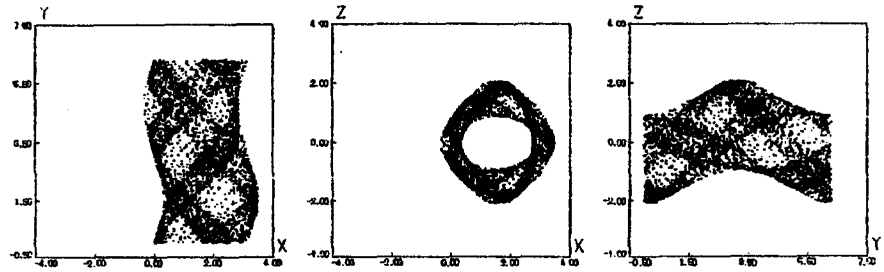


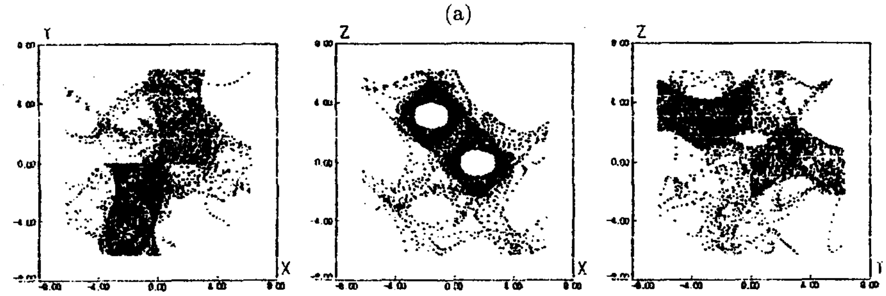
Figure 4.

and thus we know $y \geq \sqrt{2}/2$ for the numerical solutions. The projection of the trajectory on to the xy plane and yz plane, therefore, will be part of a circle and part of a ellipse respectively.

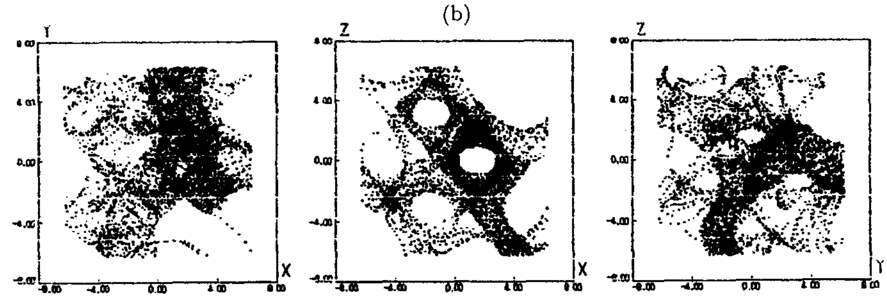
Figure 1(a) shows the phase trajectory calculated by the classical 4^{th}-order Runge-Kutta for 6000 steps when taking the step-size as 0.5. Figure 1(b) shows the projections of the trajectory on to the three planes xy, xz, yz. We see the RK method yields spiraling of phase trajectory with artificial creation of an attractor. We have also calculated out the value of $\det(J_{n+1}) = \det\left(\dfrac{\partial(x_{n+1}, y_{n+1}, z_{n+1})}{\partial(x_n, y_n, z_n)}\right)$ and the changing of the volume $\text{vol}_{n+1} = \prod_{i=1}^{n+1} \det(J_i)$ at every step. Figure 1(c) shows the value of $\det(J_i)$ at the first(left) and the last(right) 500 steps. We see the value of $\det(J_i)$ ossilates around 1.0 at first and eventually the value is always bellow 1.0. Figure 1(d) shows the changing of the volume, we see the value of vol_i decreases rapidly in the first 2000(left) steps and in the last 2000 steps(right) it's almost equal to zero.



(a)



(b)



(c)

Figure 5.

Figure 2 shows the projections of the trajectory calculated by explicit scheme (14) for 6000 steps at the same step-size. We see the trajectory is preserved except it is tortured somewhat because the explicit method is not A-stable and the step size is now big.

Though the centered Euler method is not volume-preserving in this case, the numerical result of this scheme with step-size 0.5, totally 6000 steps, is still very good as Figure 3(a) shows. The values $\det(J_i)$ and vol_i are presented in Figure 3(b) and Figure 3(c). These values remain osilliating aroud 1.0 in the first and last 500 steps, so the volume will not change greatly even in a very long time period, it always approximates, though not equal to, the true volume. Since the centered Euler method is A-stable, the phase trajectory of the system is preserved even better than the explicit method.

NUMERICAL EXPERIMENT 2. We now turn to the ABC flow [8,9]

$$\begin{cases} \dot{x} = A\sin z + C\cos y \pmod{2\pi} \\ \dot{y} = A\sin x + A\cos z \pmod{2\pi} \\ \dot{z} = C\sin y + B\cos x \pmod{2\pi}. \end{cases} \tag{17}$$

When $A = B = 1.0, C = 0.0$, (17) is integrable. In this case, the centered Euler method is volume-preserving since $\det(Df) = 0$. Projections on to the xy, xz, yz planes of the trajectories calculated by explicit method, Euler method and RK method are presented in Figure 4(a), (b), (c), respectively. The total step number is 6000, step size is 0.5 and initial values are still (1.0,1.0,1.0). We see in this case, Euler method is as good as explicit method and the RK method is still dissipative. Figure 4(d) shows the value of $\det(J_i)$ in the first and last 500 steps calculated by RK method. The value of $\det(J_i)$ oscilliates around 1.0 in the first 500 steps and almost always below 1.0 in the last 500 steps. The volume increases to 1.003 at first 2000 steps as Figure 4(e)(left) shows and then decreases slowly to 0.9 as in Figure 4(e)(right).

When $A^2 = 1, B^2 = \frac{2}{3}, C^2 = \frac{1}{3}, A > B > C > 0$, system (17) non-integrable. In this case, Euler method is non-volume-preserving. Figure 5(a) shows the results calculated by explicit method with initial values (1.0, 1.0, 1.0) and step-size 0.5. Figure 5(b) and (c) are results of Euler method and RK method with the same initial point and step-size as the explicit method, we see the result of volume-preserving explict scheme is much better than the others.

REFERENCES

1. K. Feng, Formal power series and numerical algorithms for dynamical systems, In *Proc. Intern. Conf. on Scientific Computation, Hangzhou, 1991*, (Edited by T. Chan and Z.C. Shi), pp. 28–35, World Scientific, Singapore, (1992).
2. V.I. Arnold, *Mathematical Methods of Classical Mechanics*, Springer, (1978).
3. C. Scovel, Symplectic numerical integration of Hamiltonian system, In *MSRI Workshop on the Geometry of Hamiltonian Systems*, (1989).
4. R.S. Mackay, Some aspects of the dynamical and numerics of Hamiltonian systems, In *Proceedings of the IMA Conference on "Dynamics of Numerics and Numerics of Dynamics" Bristol*, (1990).
5. K. Feng, On difference scheme and symplectic geometry, In *Proceeding of the 1984 Beijing Symposium on Differential Geomtry and Differential Equations—Computation of Partial Differential Equations*, (Edited by K. Feng), pp. 42–58, Science Press, Beijing, (1985).
6. R. Abraham and J.E. Marsden, *Foundations of Mechanics*, 2nd Edition, Benjamin/Cummings, Reading, (1980).
7. M.-Z. Qin and W.-J. Zhu, Construction of higher order symplectic schemes by composition, *Computing* **47**, 309–321 (1992).
8. T. Dombre, U. Frisch, J.M. Greene, M. Henon, A. Mehr and A.M. Soward, Chaotic Streamlines in the ABC Flows, *J. Fluid Mech.* **167**, 353–391.
9. M. Feingold, L.P. Kadanoff and O. Piro, Passive Scalars, Three-Dimensional Volume-Preserving Maps, and Chaos, *Journal of Statistical Physics* **50 (3/4)** (1988).