 CMES, vol.9, no.3, pp.255-272, 2005

Nonstandard Group-Preserving Schemes for Very Stiff Ordinary Differential Equations

Chein-Shan Liu [1]

Abstract: The group-preserving scheme developed by Liu (2001) for calculating the solutions of k-dimensional differential equations system adopted the Cayley transform to formulate the Lie group from its Lie algebra $\mathbf{A} \in so(k,1)$. In this paper we consider a more effective exponential mapping to derive $\exp(h\mathbf{A})$. In order to overcome the difficulty of numerical instabilities encountered by employing group-preserving schemes on stiff differential equations, we further combine the nonstandard finite difference method into the group-preserving schemes to obtain unconditional stable numerical methods. They provide single-step explicit time integrators for stiff differential equations. Several numerical examples are examined, some of which are compared with exact solutions showing that the nonstandard group-preserving schemes have good computational efficiency and certain accuracy.

keyword: Stiff differential equations, nonstandard group-preserving scheme, A-stable, L-stable

1 Introduction

For many systems in engineering applications, the initial value problems with stiff ordinary differential equations may occur due to the appearance of large difference of time scales exhibited in the physical models. These time scales are usually responsible for the different decaying rates of the model. Because of the speciality and complexity of these systems, the corresponding differential equations are usually called stiff differential equations, also called ill-conditioned equations. Gear method [Gear (1971)] equipped with Adams predictor–corrector method was known to be a better integrator of stiff differential equations, and is highly efficient for the solution of ill-conditioned problems for its good stability, high precision, etc. It is the first order upwind difference method and the code of Gear method has the merit of self-stability. Adams method required less computation time for its simpler iterative procedure; however, it may fail applied to strong ill-conditioned equations. When solving the ill-conditioned problems, both explicit and implicit methods will need to compute the Jacobian of the system, which may cause great inconvenience as compared to the numerical solutions of non-stiff problems. For a large system, the computation of Jacobian is often very tedious and may cause mistakes. In this paper some effective schemes are developed by considering nonstandard difference method for solving very stiff problems, which basing on the group preserving scheme proposed by Liu (2001) stated as follows.

Group-preserving scheme (GPS) can preserve the internal symmetry group of the considered system. Although we do not know previously the symmetry group of the nonlinear differential equations systems, Liu (2001) has embedded them into the augmented dynamical systems, which concern with not only the evolution of the state variables but also the evolution of the magnitude. That is, for k ordinary differential equations system:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},t), \quad \mathbf{x} \in \mathrm{R}^k, \quad t \in \mathrm{R}, \tag{1}$$

we can embed it to the following $k+1$-dimensional augemented dynamical system:

$$\frac{d}{dt}\begin{bmatrix} \mathbf{x} \\ \|\mathbf{x}\| \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{k\times k} & \frac{\mathbf{f}(\mathbf{x},t)}{\|\mathbf{x}\|} \\ \frac{\mathbf{f}^{\mathrm{T}}(\mathbf{x},t)}{\|\mathbf{x}\|} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \|\mathbf{x}\| \end{bmatrix}. \tag{2}$$

Here we assume that $\mathbf{x}$ never goes to the zero point, which leads to $\|\mathbf{x}\| > 0$ and hence the above system is well-defined.

It is obvious that the first equation in Eq. (2) is the same as the original equation (1), but the addition of the second equation gives us a Minkowskian structure of the augmented state variables of $\mathbf{X} := (\mathbf{x}, \|\mathbf{x}\|)^{\mathrm{T}}$ satisfying the cone condition:

$$\mathbf{X}^{\mathrm{T}}\mathbf{g}\mathbf{X} = 0, \tag{3}$$

[1] Department of Mechanical and Mechatronic Engineering, Taiwan Ocean University, Keelung, Taiwan, E-mail: csliu@mail.ntou.edu.tw

where

$$\mathbf{g} = \begin{bmatrix} \mathbf{I}_k & \mathbf{0}_{k\times 1} \\ \mathbf{0}_{1\times k} & -1 \end{bmatrix} \quad (4)$$

is a Minkowski metric, $\mathbf{I}_k$ is the identity matrix of order k, and the superscript T stands for the transpose. In terms of $(\mathbf{x}, \|\mathbf{x}\|)$, Eq. (3) becomes

$$\mathbf{X}^{\mathrm{T}}\mathbf{g}\mathbf{X} = \mathbf{x}\cdot\mathbf{x} - \|\mathbf{x}\|^2 = \|\mathbf{x}\|^2 - \|\mathbf{x}\|^2 = 0, \quad (5)$$

where the dot between two k-dimensional vectors denotes their Euclidean inner product. The cone condition is thus a natural constraint that we can impose on the dynamical system (2).

Consequently, we have a $k+1$-dimensional augmented system:

$$\dot{\mathbf{X}} = \mathbf{A}\mathbf{X} \quad (6)$$

with a constraint (3), where

$$\mathbf{A} := \begin{bmatrix} \mathbf{0}_{k\times k} & \frac{\mathbf{f}(\mathbf{x},t)}{\|\mathbf{x}\|} \\ \frac{\mathbf{f}^{\mathrm{T}}(\mathbf{x},t)}{\|\mathbf{x}\|} & 0 \end{bmatrix}, \quad (7)$$

satisfying

$$\mathbf{A}^{\mathrm{T}}\mathbf{g} + \mathbf{g}\mathbf{A} = \mathbf{0}, \quad (8)$$

is a Lie algebra $so(k,1)$ of the proper orthochronous Lorentz group $SO_o(k,1)$. This fact prompts us to devise the so-called group-preserving scheme, whose discretized mapping $\mathbf{G}$ exactly preserves the following properties [Liu and Chang (2004)]:

$$\mathbf{G}^{\mathrm{T}}\mathbf{g}\mathbf{G} = \mathbf{g}, \quad (9)$$

$$\det\mathbf{G} = 1, \quad (10)$$

$$G_0^0 > 0, \quad (11)$$

where G_0^0 is the 00-th component of $\mathbf{G}$. Such $\mathbf{G}$ is a proper orthochronous Lorentz group denoted by $SO_o(k,1)$. The term orthochronous used in the special relativity theory is referred to the preservation of time orientation. However, it should be understood here as the preservation of the sign of $\|\mathbf{x}\|$.

Remarkably, the original k-dimensional dynamical system (1) in E^k can be embedded naturally into an augmented $k+1$-dimensional dynamical system (6) in M^{k+1}. Although the dimension of the new system is raising one more, it has been shown that under the Lipschitz condition of

$$\|\mathbf{f}(\mathbf{x},t) - \mathbf{f}(\mathbf{y},t)\| \le L\,\|\mathbf{x}-\mathbf{y}\|, \ \ \forall\ (\mathbf{x},t),(\mathbf{y},t) \in \mathrm{D}, \quad (12)$$

where D is a domain of $\mathrm{R}^k \times \mathrm{R}$, and L is known as a Lipschitz constant, the new system has the advantage of devising group-preserving numerical scheme as follows [Liu (2001)]:

$$\mathbf{X}_{n+1} = \mathbf{G}(n)\mathbf{X}_n, \quad (13)$$

where $\mathbf{X}_n$ denotes the numerical value of $\mathbf{X}$ at the discrete time t_n, and $\mathbf{G}(n) \in SO_o(k,1)$ is the group value at time t_n.

According to the definition made by Shampine and Gear (1979), the initial value problem (1) is said to be stiff in a time interval of $t \in [a,b]$, if (i) the solution $\mathbf{x}(t)$ is slowly varying in the time interval of $[a,b]$, and if (ii) for every point $(t,\mathbf{x}(t))$ of the solution curve in the time interval of $[a,b]$ the Jacobian matrix

$$\mathbf{J} = \frac{\partial\mathbf{f}(\mathbf{x},t)}{\partial\mathbf{x}} \ \in \mathrm{R}^{k\times k} \quad (14)$$

has at least one eigenvalue whose real part is large negative, whilst the real parts of the other eigenvalues do not take large positive values.

If Eq. (1) is stiff, then the augmented Eq. (6) is also stiff. This assertion follows from the following Jacobian matrix for system (6):

$$\mathbf{J}_{\mathrm{aug}} = \begin{bmatrix} \mathbf{J} & \mathbf{0}_{k\times 1} \\ \frac{1}{\|\mathbf{x}\|}\left(\frac{\partial(\mathbf{x}\cdot\mathbf{f}(\mathbf{x},t))}{\partial\mathbf{x}}\right)^{\mathrm{T}} & \frac{-\mathbf{x}\cdot\mathbf{f}(\mathbf{x},t)}{\|\mathbf{x}\|^2} \end{bmatrix} \ \in \mathrm{R}^{(k+1)\times(k+1)}, \quad (15)$$

which has k eigenvalues $\{\lambda_1,\lambda_2,\ldots,\lambda_k\}$ as that for $\mathbf{J}$, as well as another one denoted by $\lambda_0 = -\mathbf{x}\cdot\mathbf{f}(\mathbf{x},t)/\|\mathbf{x}\|^2$. In the literature the term stiff has been used by various authors with quite different meanings. However, Spijker (1996) has reviewed various aspects of stiffness in the numerical solutions for stiff ordinary differential equations (SODEs), and has given an intutive definition of *stiff situation*, which in terms of stepsize h and the Lipschitz constant L is $hL \gg 1$.

In this paper we attempt to develop nonstandard group-preserving schemes for SODEs. It is an extension of the work of Liu (2001) by taking the stiffness of differential equations into account. Numerical schemes adopted for stiff differential equations are usually implicit. The explicit schemes that have been applied to solving the stiff problems are apparently not very effective up to now. So conventionally, the implicit schemes with variable order and stepsize are used for SODEs, among which the backward difference type is one of the most famous; see, *e.g.*, Stabrowski (1997) and references theirin. The main advantage of the implicit scheme is its stability in the long-term calculations, which is usually absent for explicit scheme. However, the main disadvantage of implicit scheme is its time consumption spent to solve the resulting nonlinear algebraic equations step-by-step.

This paper will develop highly effective schemes by an extension of the group-preserving scheme developed by Liu (2001). The new method provides an explicit single-step algorithm, and renders a more compendious numerical implementation than other conventional schemes to solve SODEs. It would be found in this study that the new method greatly reduces the computation time that is important for conducting a long-term simulation.

2 GPS for differential equations system

2.1 The Cayley transform

The group generated from $\mathbf{A} \in so(k,1)$ is known as a proper orthochronous Lorentz group, one of which is the Cayley transform

$$\mathrm{Cay}(\tau\mathbf{A}) = (\mathbf{I}-\tau\mathbf{A})^{-1}(\mathbf{I}+\tau\mathbf{A}), \tag{16}$$

a mapping from $\mathbf{A}$ to an element of $SO_o(k,1)$ for $\tau \in \mathrm{R}$ and $\tau^2 < \|\mathbf{x}\|^2/\|\mathbf{f}\|^2$. Substituting Eq. (7) for $\mathbf{A}(n)$, which denotes the value of $\mathbf{A}$ at the discrete time t_n, into the above equation yields

$$\mathrm{Cay}[\tau\mathbf{A}(n)] = \begin{bmatrix} \mathbf{I}_k + \frac{2\tau^2}{\|\mathbf{x}_n\|^2-\tau^2\|\mathbf{f}_n\|^2}\mathbf{f}_n\mathbf{f}_n^{\mathrm{T}} & \frac{2\tau\|\mathbf{x}_n\|}{\|\mathbf{x}_n\|^2-\tau^2\|\mathbf{f}_n\|^2}\mathbf{f}_n \\ \frac{2\tau\|\mathbf{x}_n\|}{\|\mathbf{x}_n\|^2-\tau^2\|\mathbf{f}_n\|^2}\mathbf{f}_n^{\mathrm{T}} & \frac{\|\mathbf{x}_n\|^2+\tau^2\|\mathbf{f}_n\|^2}{\|\mathbf{x}_n\|^2-\tau^2\|\mathbf{f}_n\|^2} \end{bmatrix}. \tag{17}$$

Inserting the above $\mathrm{Cay}[\tau\mathbf{A}(n)]$ for $\mathbf{G}(n)$ into Eq. (13) and taking its first row, we obtain

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \frac{\|\mathbf{x}_n\|^2+\tau\mathbf{f}_n\cdot\mathbf{x}_n}{\|\mathbf{x}_n\|^2-\tau^2\|\mathbf{f}_n\|^2}h\mathbf{f}_n = \mathbf{x}_n + \eta_n\mathbf{f}_n. \tag{18}$$

In the above $\mathbf{x}_n$ denotes the numerical value of $\mathbf{x}$ at the discrete time t_n, τ is one half of the time increment, *i.e.*, $\tau := h/2$, $\mathbf{f}_n$ denotes $\mathbf{f}(\mathbf{x}_n,t_n)$, and η_n is an adaptive factor. In order to meet the property (11), we require the stepsize of scheme (18) being constrained by $h < 2\|\mathbf{x}_n\|/\|\mathbf{f}_n\|$. Under this condition we have

$$h < \frac{2\|\mathbf{x}_n\|}{\|\mathbf{f}_n\|} \Longleftrightarrow G_0^0 > 0 \Longrightarrow \eta_n > 0. \tag{19}$$

Some properties of preserving the fixed point behavior of the above numerical scheme (18) have been investigated by Liu (2001), and applying it to non-stiff differential equations has revealed that it is easy to implement numerically and has high computational efficiency and accuracy as discussed by Liu (2001). However, for stiff differential equations the stepsize may be constrained very small in order to meet the requirement in Eq. (19). In Section 3 we will propose some modifications.

2.2 Exponential mapping

An exponential mapping of $\mathbf{A}(n)$ admits a closed-form representation:

$$\exp[h\mathbf{A}(n)] = \begin{bmatrix} \mathbf{I}_k + \frac{(a_n-1)}{\|\mathbf{f}_n\|^2}\mathbf{f}_n\mathbf{f}_n^{\mathrm{T}} & \frac{b_n\mathbf{f}_n}{\|\mathbf{f}_n\|} \\ \frac{b_n\mathbf{f}_n^{\mathrm{T}}}{\|\mathbf{f}_n\|} & a_n \end{bmatrix}, \tag{20}$$

where

$$a_n := \cosh\left(\frac{h\|\mathbf{f}_n\|}{\|\mathbf{x}_n\|}\right), \quad b_n := \sinh\left(\frac{h\|\mathbf{f}_n\|}{\|\mathbf{x}_n\|}\right). \tag{21}$$

Substituting the above $\exp[h\mathbf{A}(n)]$ for $\mathbf{G}(n)$ into Eq. (13) and taking its first row, we obtain

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \eta_n\mathbf{f}_n, \tag{22}$$

where the adaptive factor

$$\eta_n := \frac{b_n\|\mathbf{x}_n\|\|\mathbf{f}_n\| + (a_n-1)\mathbf{f}_n\cdot\mathbf{x}_n}{\|\mathbf{f}_n\|^2} \tag{23}$$

is varying step-by-step. From $a_n > 1$, $\forall h > 0$ and $\|\mathbf{f}_n\|\|\mathbf{x}_n\| \geq \mathbf{f}_n\cdot\mathbf{x}_n \geq -\|\mathbf{f}_n\|\|\mathbf{x}_n\|$, we can prove that

$$\frac{\left[\exp\left(\frac{h\|\mathbf{f}_n\|}{\|\mathbf{x}_n\|}\right)-1\right]\|\mathbf{x}_n\|}{\|\mathbf{f}_n\|} \geq \eta_n \geq \frac{\left[1-\exp\left(-\frac{h\|\mathbf{f}_n\|}{\|\mathbf{x}_n\|}\right)\right]\|\mathbf{x}_n\|}{\|\mathbf{f}_n\|} > 0, \quad \forall h > 0. \tag{24}$$

Although this scheme is group properties preserved for all $h > 0$, and does not endure the same shortcoming as the one for scheme (18), the factor $h\|\mathbf{f}_n\|/\|\mathbf{x}_n\|$ in Eq. (21) may render scheme (22) blow-up when apply it to stiff differential equations by using an h not small enough.

2.3 Translation in state space

At the very beginning we have assumed that $\mathbf{x}$ never passes through the zero point, which means that $\|\mathbf{x}\| > 0$ and hence schemes (18) and (22) are workable. However, we may face such a system whose orbit $\mathbf{x}$ may pass the zero point or may tend to it. For this case we should modify the above methods by simply considering the translation of $\mathbf{x}$ as to be shown below.

The property (24) holds for numerical scheme (22) with any stepsize $h > 0$. So we may choose large h in the calculation without leading to inconsistent numerical result by scheme (22) for non-stiff differential equations. However, this is not the case for scheme (18). For a large h with $h > 2\|\mathbf{x}_n\|/\|\mathbf{f}_n\|$, the adaptive factor in Eq. (18) is negative, and hence scheme (18) may give improperly qualitative behavior of numerical solution.

If $\mathbf{x}$ runs near the zero point, in order to meet the requirement of $\|\mathbf{x}\| > \tau\|\mathbf{f}\|$, it not only needs τ very small but also leads to a much stringent constraint on the choice of a suitable stepsize in the calculation. In practice, we may encounter the problem of dividing by a very small denominator in Eq. (18) when fixing the stepsize, and such division always renders the scheme failing to calculate the solution.

In order to avoid the above difficulties of possibly dividing by a very small denominator or even dividing by a zero number when $\|\mathbf{x}\| = 0$, let us translate the state variable $\mathbf{x}$ to a new variable $\mathbf{u}$ by a constant vector $\mathbf{b}$:

$$\mathbf{u} = \mathbf{x} + \mathbf{b}, \tag{25}$$

such that $\mathbf{u}$ is far apart from the zero point. Therefore, we have a new system of ODEs given by

$$\dot{\mathbf{u}} = \mathbf{F}(\mathbf{u}, t) := \mathbf{f}(\mathbf{u} - \mathbf{b}, t). \tag{26}$$

Although $\mathbf{x}$ may close to or even go to the zero point we do not worry about the zeroness of $\|\mathbf{u}\|$ because of $\|\mathbf{u}\| = \|\mathbf{b}\| > 0$ when $\mathbf{x} = \mathbf{0}$. Even Lee, Chen and Hung (2002) have applied the above technique to treat some differential equations with mild stiffness, however, we should note that the translation (25) does not remove the stiffness of the considered system because $\partial\mathbf{F}/\partial\mathbf{u} = \partial\mathbf{f}/\partial\mathbf{x}$ has the same eigenvalues.

Now, replacing $\mathbf{x}$ by $\mathbf{u}$ and $\mathbf{f}$ by $\mathbf{F}$ in Eq. (18) we obtain a similar scheme for the new system (26) as follows:

$$\mathbf{u}_{n+1} = \mathbf{u}_n + \frac{\|\mathbf{u}_n\|^2 + \tau\mathbf{F}_n \cdot \mathbf{u}_n}{\|\mathbf{u}_n\|^2 - \tau^2\|\mathbf{F}_n\|^2} h\mathbf{F}_n. \tag{27}$$

This yields a time marching scheme for $\mathbf{u}$ and thus gives the solution of $\mathbf{x}$ by calculating $\mathbf{x} = \mathbf{u} - \mathbf{b}$ at each time step.

By the same token, scheme (22) can be applied to the new variable $\mathbf{u}$ with the following formula:

$$\mathbf{u}_{n+1} = \mathbf{u}_n + \frac{b_n\|\mathbf{u}_n\|\|\mathbf{F}_n\| + (a_n - 1)\mathbf{F}_n \cdot \mathbf{u}_n}{\|\mathbf{F}_n\|^2}\mathbf{F}_n, \tag{28}$$

where

$$a_n := \cosh\left(\frac{h\|\mathbf{F}_n\|}{\|\mathbf{u}_n\|}\right), \quad b_n := \sinh\left(\frac{h\|\mathbf{F}_n\|}{\|\mathbf{u}_n\|}\right). \tag{29}$$

Some numerical examples will be shown later that such strategy is effective for certain mildly stiff differential equations, but it still fails to calculate the solutions for very stiff differential equations. Basically, by merely translating the state vector it can not lighten the high stiffness of differential equations. It means that Eq. (26) is still highly stiff if the original Eq. (1) is. Under this very stringent condition we should consider another effective method as follows.

3 Stable group-preserving scheme for stiff differential equations

The main idea of nonstandard finite difference [e.g., Mickens (1994, 1999), Mickens and Ramadhani (1994) and Mickens and Smith (1990)] is replacing the Euler approximation of $\dot{\mathbf{X}}$

$$\dot{\mathbf{X}} \longrightarrow \frac{\mathbf{X}_{n+1} - \mathbf{X}_n}{h}, \tag{30}$$

by a nonstandard approximation

$$\dot{\mathbf{X}} \longrightarrow \frac{\mathbf{X}_{n+1} - \mathbf{X}_n}{\phi(h)}, \tag{31}$$

where $\phi(h)$ is called a denominator function having the properties of $\phi(h) > 0$ and $\phi(h) = h + O(h^2)$.

For stiff differential equations we let

$$\phi(h) := \frac{1 - \exp(-Lh)}{L}, \tag{32}$$

where L is the Lipschitz constant of Eq. (1). If the considered system (1) is stiff in some time interval of $t \in [a, b]$, and its Jacobian matrix $\mathbf{J}$ is continuous and bounded in $[a, b]$, then we let the Lipschitz constant to be

$$L = \|\mathbf{J}\| \geq \max\{|\lambda_i| : i = 1, 2, \ldots, k\}, \tag{33}$$

where $\|\mathbf{J}\|$ stands for the Euclidean norm of $\mathbf{J}$. Hence, the Lipschitz constant is a large value for stiff differential equations.

The replacement of h by $\phi(h)$ in Eq. (31) inspires us to replace the h's in Eqs. (18) and (22) by $\phi(h)$; consequently, we have

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \frac{4\|\mathbf{x}_n\|^2 + 2\phi \mathbf{f}_n \cdot \mathbf{x}_n}{4\|\mathbf{x}_n\|^2 - \phi^2\|\mathbf{f}_n\|^2}\phi \mathbf{f}_n, \tag{34}$$

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \frac{b_n\|\mathbf{x}_n\|\|\mathbf{f}_n\| + (a_n - 1)\mathbf{f}_n \cdot \mathbf{x}_n}{\|\mathbf{f}_n\|^2}\mathbf{f}_n, \tag{35}$$

where

$$a_n := \cosh\left(\frac{\phi\|\mathbf{f}_n\|}{\|\mathbf{x}_n\|}\right), \;\; b_n := \sinh\left(\frac{\phi\|\mathbf{f}_n\|}{\|\mathbf{x}_n\|}\right). \tag{36}$$

Now we investigate what advantages can be gained by schemes (34) and (35). From Eq. (32) and $\|\mathbf{f}_n\| \le L\,\|\mathbf{x}_n\|$ it follows that

$$\frac{\phi\|\mathbf{f}_n\|}{\|\mathbf{x}_n\|} \le \phi L < 1, \;\; \forall h > 0. \tag{37}$$

Hence, the denominator in Eq. (34) is positive, *i.e.*,

$$4\|\mathbf{x}_n\|^2 - \phi^2\|\mathbf{f}_n\|^2 = \|\mathbf{x}_n\|^2\left(4 - \frac{\phi^2\|\mathbf{f}_n\|^2}{\|\mathbf{x}_n\|^2}\right) > 0. \tag{38}$$

It guarantees that the adaptive factor in Eq. (34) is always positive, that is,

$$\eta_n := \frac{4\|\mathbf{x}_n\|^2 + 2\phi \mathbf{f}_n \cdot \mathbf{x}_n}{4\|\mathbf{x}_n\|^2 - \phi^2\|\mathbf{f}_n\|^2}\phi > 0, \;\; \forall h > 0. \tag{39}$$

Similarly, due to Eq. (37) the coefficients a_n and b_n in Eq. (36) are bounded, and hence the over-flow which may happen for scheme (22) can be avoided. The combination of nonstandard method with group-preserving schemes renders the new numerical schemes (34) and (35) always stable. More practically, they are unconditionally stable. This result is very important for stiff differential equations, because the dominant factor to choose a suitable stepsize for stiff differential equation is its stability, not its accuracy, as shown by Shampine and Gear (1979). However, we can see in the later that schemes (34) and (35) still have an accuracy in the order of $O(h)$.

Furthermore, scheme (34) preserves the fixed point and the property of the original differential equations system. Under the above condition (39), it is obvious that

$$\mathbf{x}_{n+1} = \mathbf{x}_n \Longleftrightarrow \mathbf{f}_n = \mathbf{0}. \tag{40}$$

This means that $\mathbf{x}_n$ is a fixed point of the discretized mapping (34) if and only if the point $\mathbf{x}$ is an equilibrium (critical, fixed) point of the system (1).

We next investigate the property of the fixed point. The Jacobian of the mapping (34) is

$$J := \frac{\partial \mathbf{x}_{n+1}}{\partial \mathbf{x}_n} = \mathbf{I}_k + \mathbf{f}_n\left(\frac{\partial \eta_n}{\partial \mathbf{x}_n}\right)^{\mathrm{T}} + \eta_n \mathbf{J}_n, \tag{41}$$

where $\mathbf{J}_n$ denotes the value of $\mathbf{J}$ at time $t = t_n$. At the fixed point $\mathbf{f}_n = \mathbf{0}$, we have $\eta_n = \phi$, and thus

$$J = \mathbf{I}_k + \phi \mathbf{J}_n. \tag{42}$$

Obviously, J has eigenvalues $\{1+\phi\lambda_i,\; i = 1,2,\ldots,k\}$. For the eigenvalue $\lambda = \mathrm{Re}(\lambda) + i\mathrm{Im}(\lambda)$ of $\mathbf{J}$, where the prefixes Re and Im denote respectively the real and imaginary parts, we let $|\lambda|^2 = (\mathrm{Re}(\lambda))^2 + (\mathrm{Im}(\lambda))^2$ and we can prove the following implications:

$$\mathrm{Re}(\lambda) > 0 \Longrightarrow |1+\phi\lambda| > 1, \; \forall h > 0, \tag{43}$$

$$\mathrm{Re}(\lambda) = 0 \Longrightarrow |1+\phi\lambda| \ge 1, \; \forall h > 0, \tag{44}$$

$$\mathrm{Re}(\lambda) < 0 \Longrightarrow \begin{cases} |1+\phi\lambda| < 1, \; \forall h > 0 \\ \quad \text{if } L \ge \frac{|\lambda|^2}{-2\mathrm{Re}(\lambda)}, \\ |1+\phi\lambda| < 1, \\ \quad h < \ln\left(1 + \frac{2L\,\mathrm{Re}(\lambda)}{|\lambda|^2}\right)^{-1/L} \\ \quad \text{if } L < \frac{|\lambda|^2}{-2\mathrm{Re}(\lambda)}. \end{cases} \tag{45}$$

From the following equation

$$|1+\phi\lambda| = \sqrt{1 + 2\phi\mathrm{Re}(\lambda) + \phi^2[(\mathrm{Re}(\lambda))^2 + (\mathrm{Im}(\lambda))^2]}, \tag{46}$$

and $\phi > 0$, Eq. (43) follows directly. For the second case, substituting $\mathrm{Re}(\lambda) = 0$ into Eq. (46) leads to Eq. (44), of which the equality holds only under the condition of $\mathrm{Im}(\lambda) = 0$. Now we prove the third implication (45). If we let $|1+\phi\lambda| = 1$, by Eq. (46) we obtain

$$\phi\left(2\mathrm{Re}(\lambda) + \phi[(\mathrm{Re}(\lambda))^2 + (\mathrm{Im}(\lambda))^2]\right) = 0. \tag{47}$$

Because of $\phi > 0$, it follows that

$$\phi = \frac{-2\mathrm{Re}(\lambda)}{|\lambda|^2}, \tag{48}$$

which together with Eq. (32) leads to

$$1 + \frac{2L\,\mathrm{Re}(\lambda)}{|\lambda|^2} = \exp(-Lh). \tag{49}$$

The right-hand side is positive. Under the condition of $1+2L\,\mathrm{Re}(\lambda)/|\lambda|^2>0$, the above equation has a solution for h:

$$h=\ln\left(1+\frac{2L\,\mathrm{Re}(\lambda)}{|\lambda|^2}\right)^{-1/L}. \tag{50}$$

This completes the proof of Eq. (45).

From Eq. (43) it follows that the property of unstable fixed point is not changed by the mapping (34). From Eq. (44) it follows that the fixed point of neutral type is not preserved by the mapping (34). Finally, from Eq. (45) we know that the mapping (34) preserves the property of stable fixed point for all $h>0$ for the most cases, unless the eigenvalue at the fixed point has very small negative real part.

4 Numerical examples

4.1 Example 1

Consider the following planar dynamical system:

$$\dot{x}_1=-x_1+\frac{2x_2}{\ln(x_1^2+x_2^2)},\ \ \dot{x}_2=-x_2-\frac{2x_1}{\ln(x_1^2+x_2^2)}, \tag{51}$$

whose solution, in terms of the polar coordinates (r,θ), can be expressed as

$$r(t)=r_0e^{-t},\ \ \theta(t)=\theta_0+\ln\left(1-\frac{t}{\ln r_0}\right),$$

where $r_0=r(0)$ and $\theta_0=\theta(0)$ are initial values. This example is not of stiff type; however, we use it to demonstrate the difference of schemes (18) and (22).

Obviously, $(\bar{x}_1,\bar{x}_2)=(0,0)$ is a stable fixed point. Letting $r_0=10$ and $\theta_0=\pi/6$ and applying schemes (18) and (22) to Eq. (51) with $h=0.001$ sec in the time interval of $0\le t\le 2$ sec, indicates that the accuracy of both schemes is in the order of $O(h)$ as shown in Fig. 1(a). However, with $r_0=20000$ and $h=2$ sec, scheme (18) producing a solution sticks on the initial point and fails to converge to the stable fixed point, but scheme (22) still producing a solution converges to the correct stable fixed point as shown in Fig. 1(b).

4.2 Example 2

Consider the following two-dimensional SODEs:

$$\dot{x}_1=9x_1+24x_2+5\cos t-\frac{1}{3}\sin t,\ \ x_1(0)=\frac{4}{3}, \tag{52}$$

$$\dot{x}_2=-24x_1-51x_2-9\cos t+\frac{1}{3}\sin t,\ \ x_2(0)=\frac{2}{3}, \tag{53}$$

whose solutions are given by

$$x_1(t)=2e^{-3t}-e^{-39t}+\frac{1}{3}\cos t,$$

$$x_2(t)=-e^{-3t}+2e^{-39t}-\frac{1}{3}\cos t.$$

The exact solutions show that after a long time the orbit will pass through near the zero point, and tend to a straight line with slope -1 in the plane (x_1,x_2).

We first apply schemes (18) and (22) to Eqs. (52) and (53) without considering translation. The results are displayed in Fig. 2, where $h=0.001$ sec was used. Scheme (18) gives incorrect numerical solution as shown by the time history of x_1, whose adaptive factor is negative at some time moments and zero in some time intervals. This indicates that under such $h=0.001$ sec scheme (18) is unstable, but scheme (22) is still stable. The latter gives correct result as shown in Fig. 2(a), whose error as shown in Fig. 2(b) is within the order of $O(h)$. We plot the values of $\sqrt{x_1^2+x_2^2}$ in Fig. 2(c). It can be seen that when $\sqrt{x_1^2+x_2^2}$ approaches to zero the adaptive factor of scheme (22) as shown in Fig. 2(d) is irregular appearing spikes, and at the same time the error as shown in Fig. 2(b) has peaks. So we applied schemes (27) and (28) to the above system by considering $(b_1,b_2)=(1,1)$ and $h=0.001$ sec. They lead to the same results, and both give correct numerical solutions as shown in Fig. 3. It is clear that by translating the state vector with the above (b_1,b_2) we obtain more smooth curves of adaptive factors as shown in Figs. 3(d) and 3(e), and the peaks of the error as shown in Fig. 2(b) disappear in Figs. 3(b) and 3(c). Finally, when the stepsize was increased to $h=0.05$ sec, and kept the above translation unchanged, Fig. 4(a) displays the trajectories of the numerical solutions in the plane (x_1,x_2), which converge to the steady state solution very soon, even they have inconsistent transient behaviors. In the transient stage both adaptive factors as shown in Figs. 4(d) and 4(e) are highly oscillatory. This example shows that the translation of state vector really gives improvement of the numerical solutions.

4.3 Example 3

Consider a very stiff system of nonlinear three-dimensional ODEs:

$$\dot{x}_1=-0.04x_1+10^4x_2x_3,\ \ x_1(0)=1, \tag{54}$$

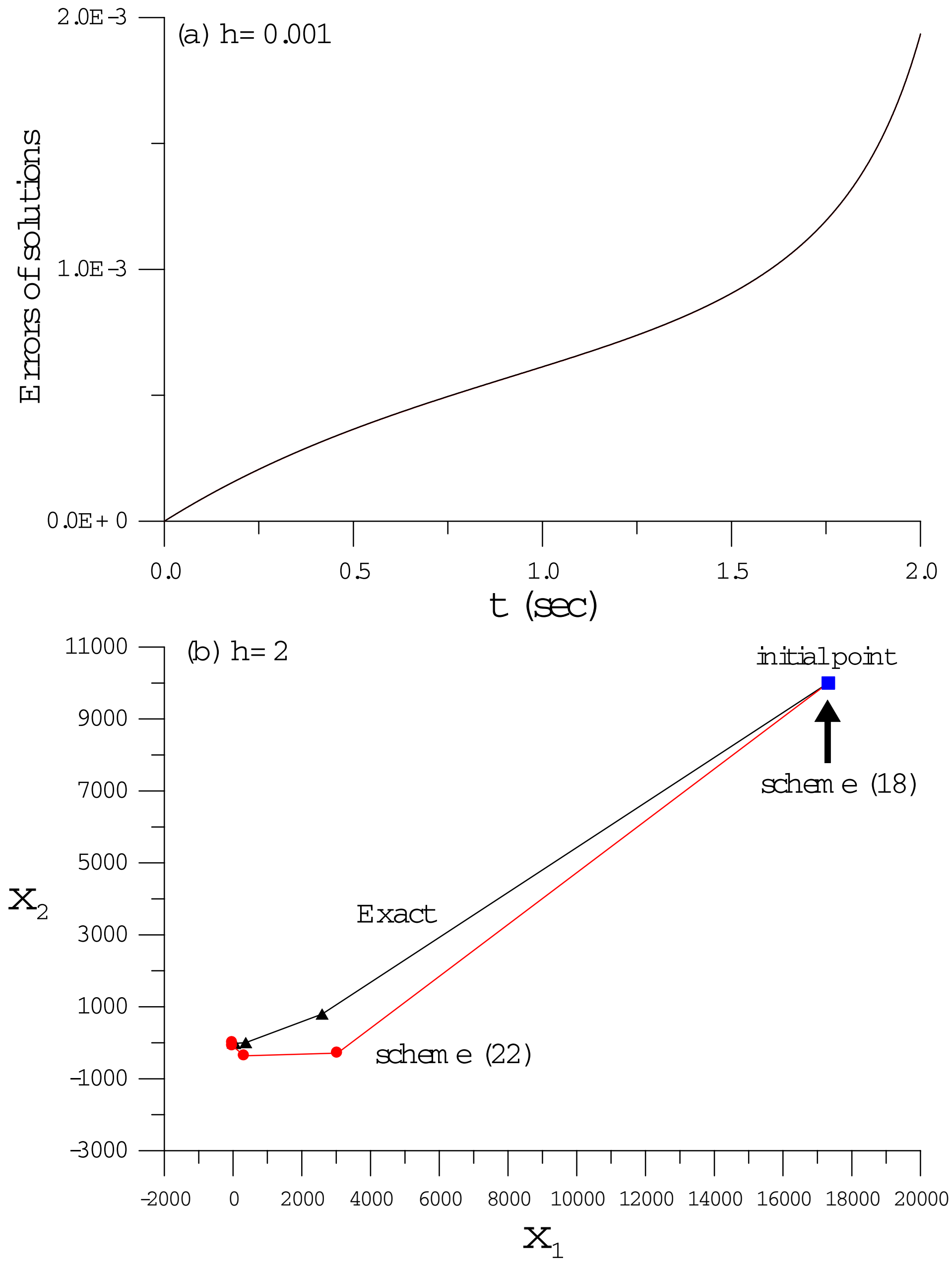

Figure 1 : The numerical solutions of Example 1 were obtained by schemes (18) and (22). For small stepsize with $h = 0.001$ sec both schemes are better coincident with exact solution, but for large stepsize with $h = 2$ sec only scheme (22) gives appropriate solution.

$$\dot{x}_2 = 0.04x_1 - 10^4 x_2 x_3 - 3\times 10^7 x_2^2, \quad x_2(0) = 0, \tag{55}$$

$$\dot{x}_3 = 3\times 10^7 x_2^2, \quad x_3(0) = 0. \tag{56}$$

The constant coefficients range from 10^{-2} to 10^7 over nine orders of magnitude. It is the source of the stiffness of this problem. The above system has been discussed by Robertson (1996), Lapidus and Seinfeld (1971), and Stabrowski (1997). First we note that $x_1 + x_2 + x_3 - 1 = 0$ is an invariant of Eqs. (54)-(56), and we may question that does the numerical scheme preserve this invariant?

Applying schemes (18) and (22) to the above system with $h = 0.0003$ sec, we show the errors of invariant in

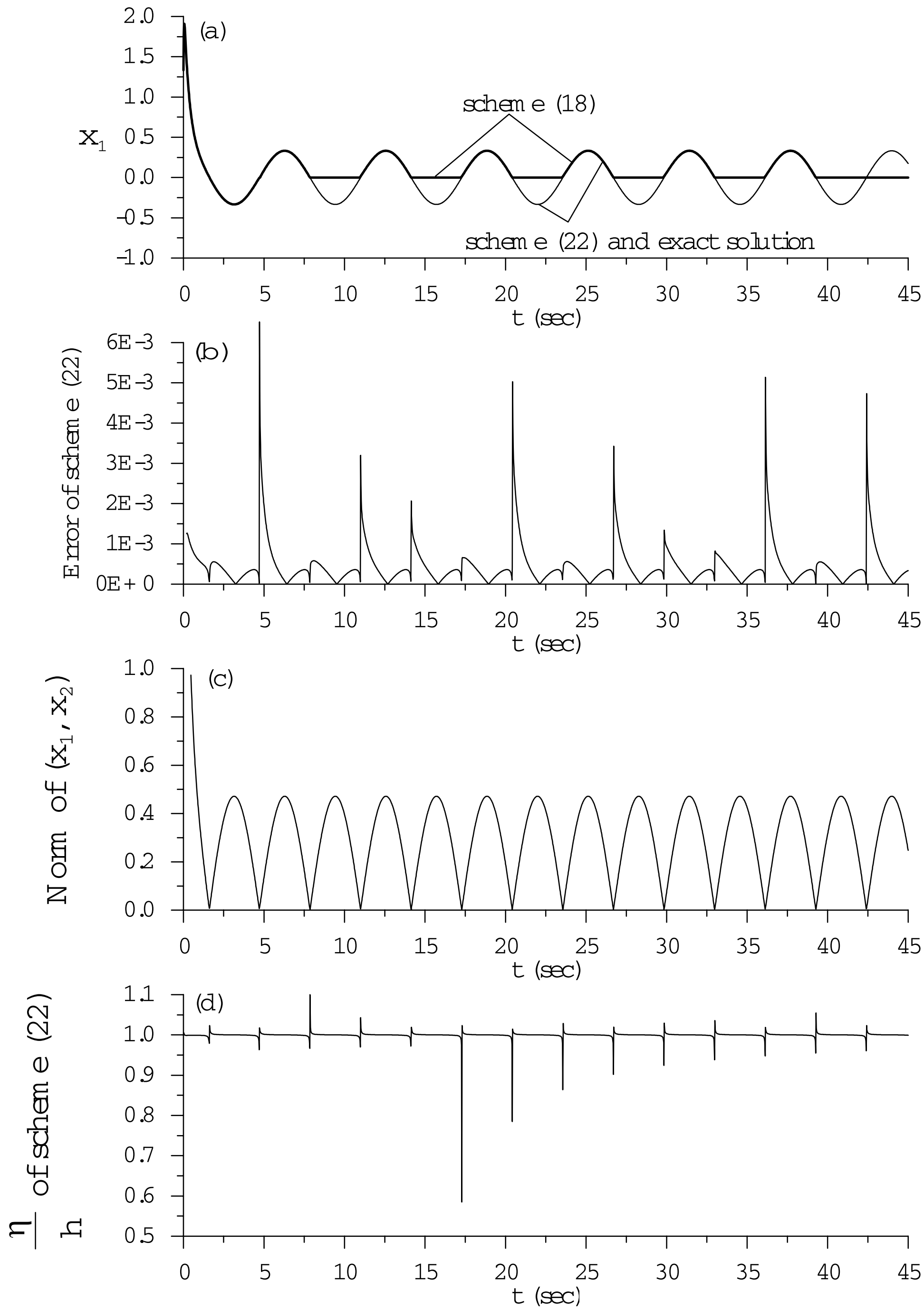

Figure 2 : The numerical solutions of Example 2 were obtained by schemes (18) and (22) with stepsize $h =$ 0.001 sec. Without considering translation scheme (18) gives inappropriate solution because its adaptive factor may be negative. However, scheme (22) still gives correct solution, but some error peaks appear.

Fig. 5. It can be seen that the errors are in the order of 10^{-14}. Even schemes (18) and (22) have the advantage to preserve the invariant, they still fail to calculate the long-term behavior. The time spent to approach the stable fixed point $(\bar{x}_1, \bar{x}_2, \bar{x}_3) = (0, 0, 1)$ may over 10^8 sec as shown in Fig. 6. The three eigenvalues of the Jacobian matrix are $(\lambda_1, \lambda_2, \lambda_3) = (0, 0, -0.04)$ at the initial point, but $(\lambda_1, \lambda_2, \lambda_3) = (0, 0, -10^4)$ at the fixed point.

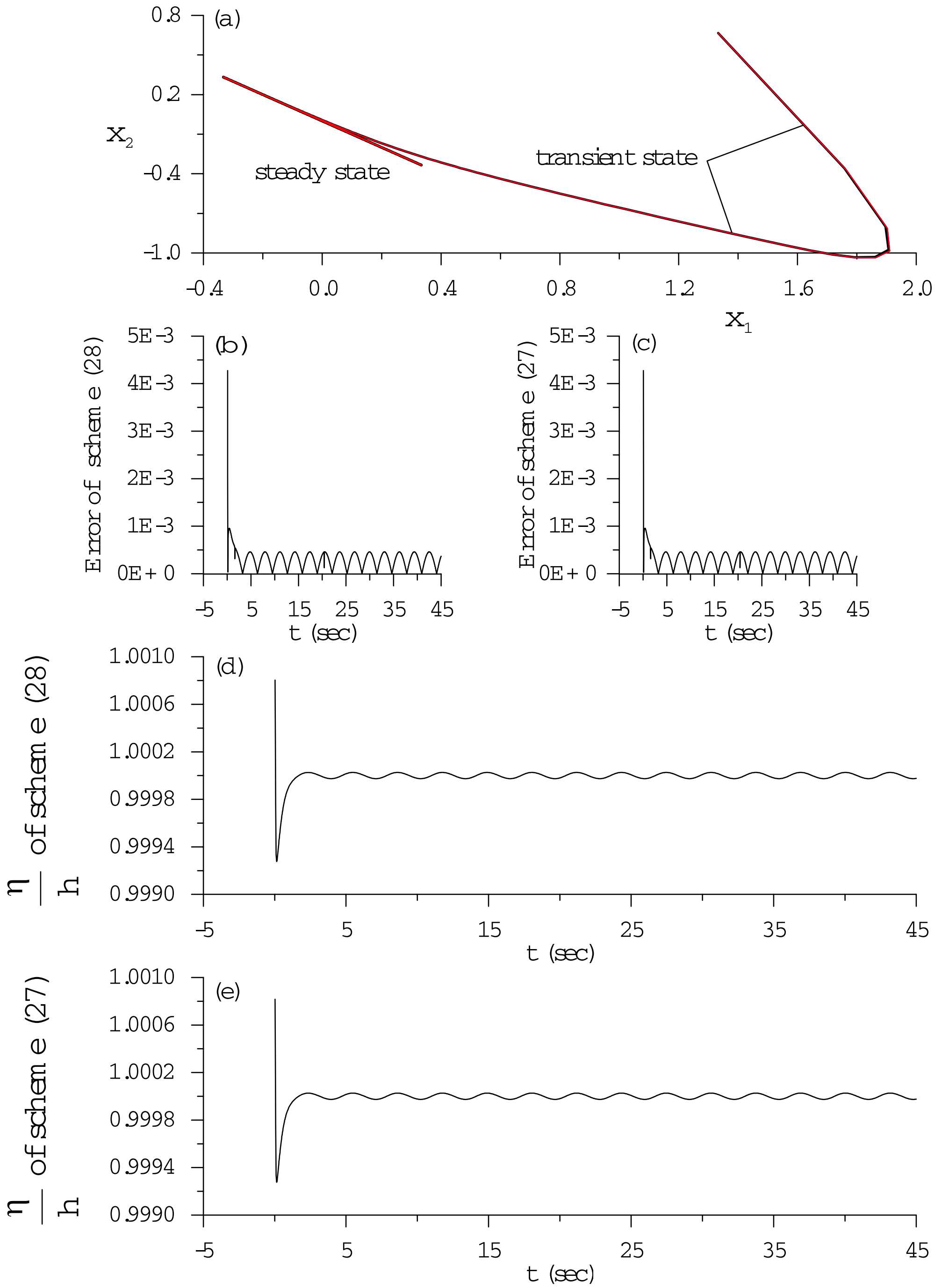

Figure 3 : The numerical solutions of Example 2 were obtained by schemes (27) and (28) with stepsize $h = 0.001$ sec and translation $(b_1, b_2) = (1, 1)$. They lead to the same results compatible with exact solution. The peaks of numerical solutions error disappear, which illustrate the effect of considering translation.

We thus apply the modified scheme (34) to the above system by letting $L = 10^4$ in Eq. (32), and hence, $\phi = [1 - \exp(-10^4 h)]/10^4$. Even the result shown in Fig. 6 used a large stepsize $h = 2$ sec, it still requires 5×10^7 steps to the final long time 10^8 sec. From Fig. 6(b) it can be seen that the three values of $(\dot{x}_1, \dot{x}_2, \dot{x}_3)$ tend to zero.

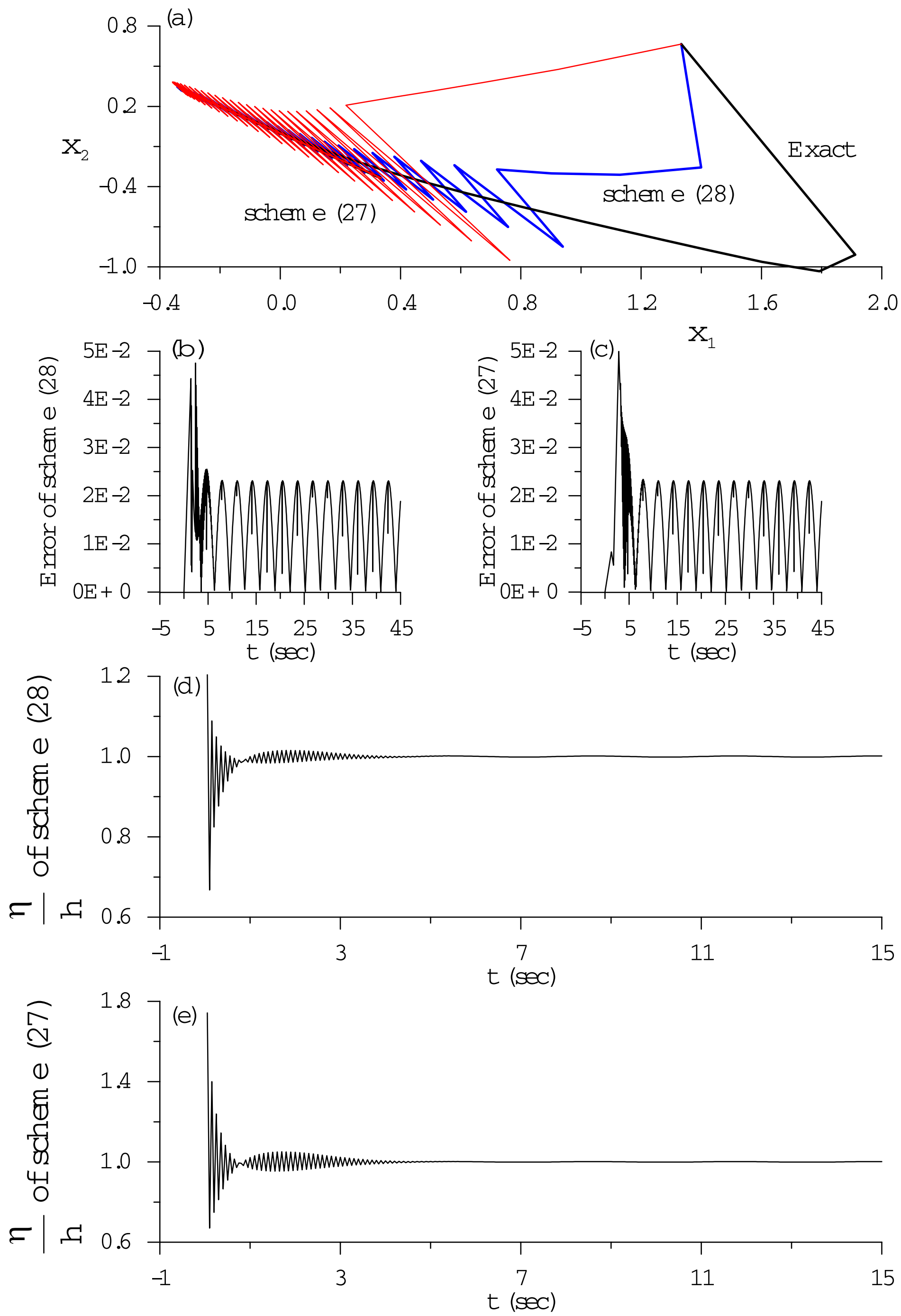

Figure 4 : The numerical solutions of Example 2 were obtained by schemes (27) and (28) with stepsize $h = 0.05$ sec and translation $(b_1, b_2) = (1, 1)$. Even they have not accurate transient solutions but tend to the exact steady state solution.

If we use schemes (18) and (22) to calculate the above long-term results, of which the stepsize $h = 0.0003$ sec is chosen for a stability consideration, it would need more than 10^{11} steps and a lot of computational time to finish the same job. Next, we use scheme (34) to calculate the solution in the range of $t < 10$ sec with $h = 0.00001$ sec.

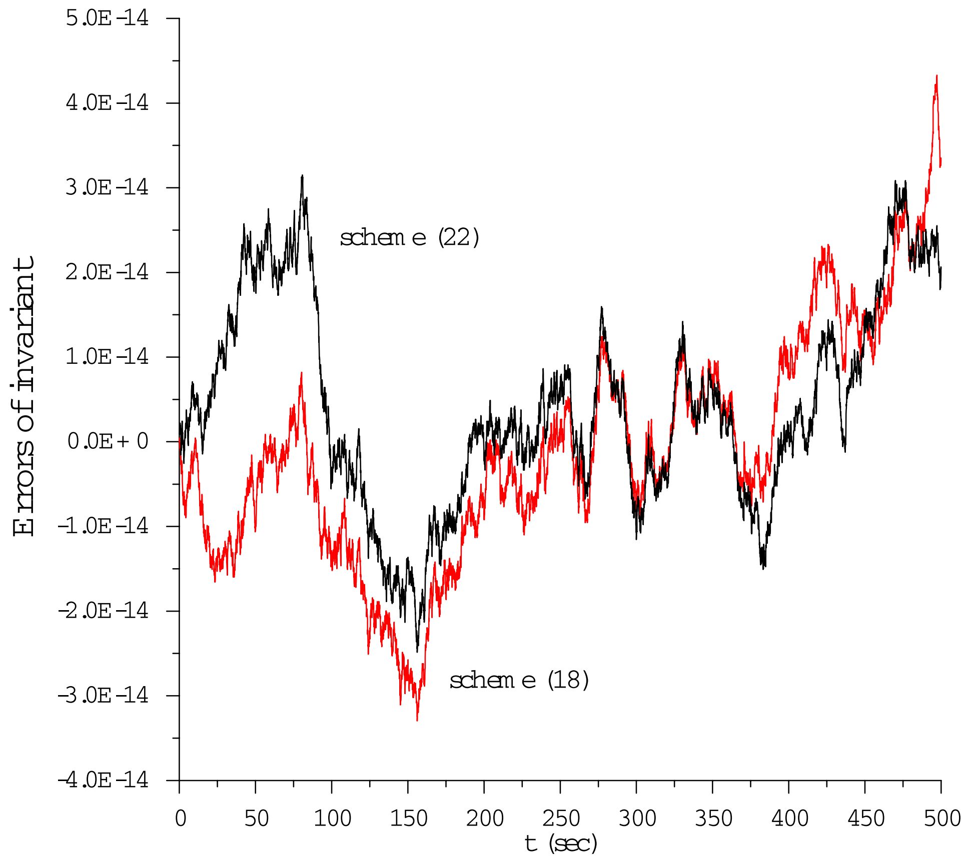

Figure 5 : The numerical solutions of Example 3 were obtained by schemes (18) and (22) with stepsize $h = 0.0003$ sec, whose errors of the invariant $x_1 + x_2 + x_3 - 1 = 0$ as can be seen are within the order of 10^{-14}.

The result as shown in Fig. 7 is compatible with that calculated by LSODE integrator as shown in Fig. 4.1 of the paper by Byrne and Hindmarsh (1987). The graphical results show how x_2 starts at 0, builds to about 3.6×10^{-4} at time $t = 2 \times 10^{-3}$ sec and decays.

4.4 Example 4

Consider another stiff system of nonlinear three-dimensional ODEs:

$$\dot{x}_1 = -0.013x_2 - 1000x_1x_2 - 2500x_1x_3, \ \ x_1(0) = 0, \quad (57)$$

$$\dot{x}_2 = -0.013x_2 - 1000x_1x_2, \ \ x_2(0) = 1, \quad (58)$$

$$\dot{x}_3 = -2500x_1x_3, \ \ x_3(0) = 1 \quad (59)$$

in the range of $0 \le t < 50$ sec. The above system has been discussed by Brunner (1974) in the problem of chemical reaction. Brunner has introduced a method of recursive collocation to solve such system. We compare the numerical value calculated from scheme (18) by using $h = 0.0001$ sec with the exact values and the calculated values from Brunner at the final time in Table 1. The error by scheme (18) as can be seen is smaller than those calculated by the Brunner method.

4.5 Example 5

We consider a linear stiff system of Rosenbrock and Storey (1966):

$$\dot{x}_1 = -1000x_1, \ \ x_1(0) = 1, \quad (60)$$

$$\dot{x}_2 = 0.909x_1 - x_2, \ \ x_2(0) = 0.999 \quad (61)$$

in the range of $0 \le t < 0.024$ sec. The exact solutions are

$$x_1(t) = \exp(-1000t),$$

$$x_2(t) = -\frac{0.909}{999}\exp(-1000t) + \frac{998.91}{999}\exp(-t).$$

In the calculations we fix the stepsize to be $h = 0.003$ sec, and apply scheme (34) to the above system by letting $L =$

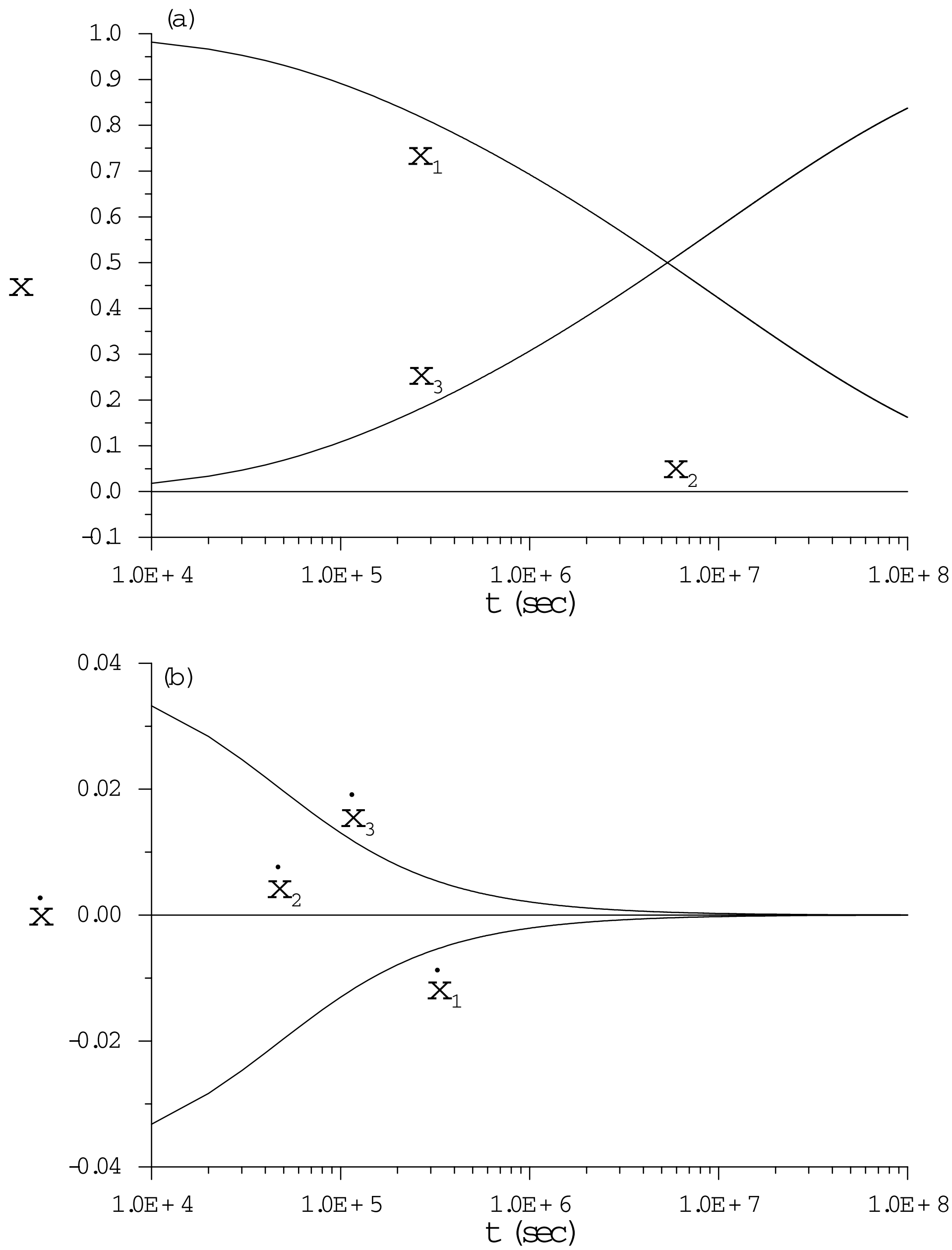

Figure 6 : The numerical solutions of Example 3 were obtained by schemes (34) and (35) with stepsize $h = 2$ sec. This example shows that it needs a huge time more than 10^8 sec to approach the stable fixed point.

10^3 in Eq. (32), and hence $\phi = [1-\exp(-10^3 h)]/10^3$. We also calculate the above system by the fourth-order Runge-Kutta method (RK4). The calculated results were compared with the exact solutions at the final time $t = 0.024$ sec in Table 2. It can be seen that the RK4 even has a certain accurate solution on the slowly changing component x_2, but it induces a larger error on the rapidly changing component x_1.

4.6 *Example 6*

Let us consider another linear stiff system given by Lapidus and Schiesser (1976):

$$\dot{x}_1 = -0.1x_1 - 49.9x_2, \quad x_1(0) = 2, \tag{62}$$

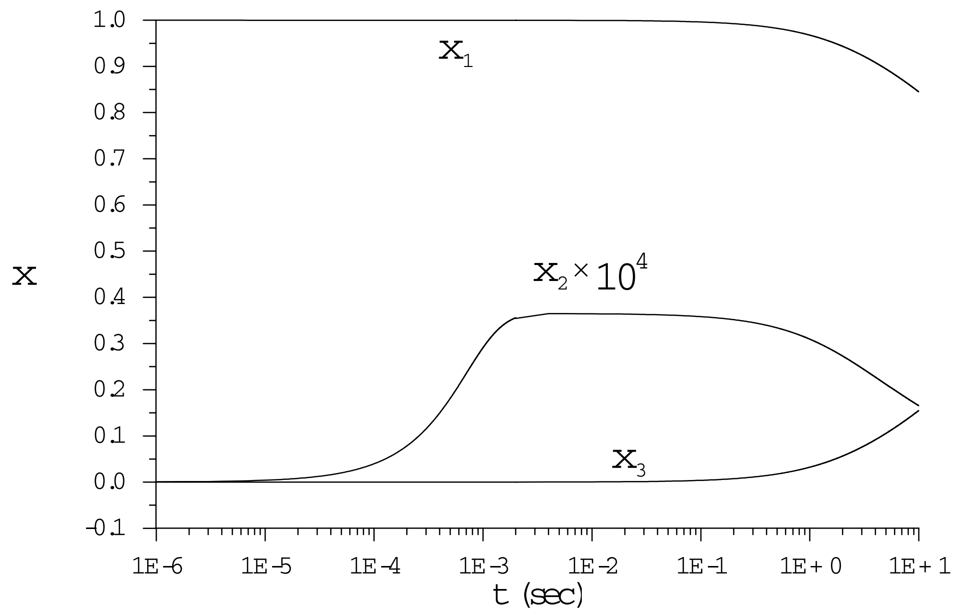

Figure 7 : The numerical solution of Example 3 was calculated in the range of $t < 10$ sec with $h = 0.00001$ sec. The result shows how x_2 starts at 0, builds to about 3.6×10^{-4} at time $t = 2 \times 10^{-3}$ sec and decays.

Table 1 : Numerical results of Example 4

$t = 50$ sec	$x_1(t)$	$x_2(t)$	$x_3(t)$
Exact	-1.89371×10^{-6}	0.5976547	1.4023434
Brunner	-1.893×10^{-6}	0.5974750	1.4025231
Error of Brunner	-0.00071×10^{-6}	0.0001797	-0.0001797
scheme (18)	-1.893386×10^{-6}	0.5976546	1.4023436
Error of scheme (18)	-0.000324×10^{-6}	0.0000001	-0.0000002

Table 2 : Numerical results of Example 5

$t = 0.024$ sec	$x_1(t)$	$x_2(t)$
Exact	$0.37751345442791 \times 10^{-10}$	0.97619772471953
RK4	$0.12776784956455 \times 10^{2}$	0.96457203308391
Error of RK4	$0.12776784956417 \times 10^{2}$	$0.11625691635613 \times 10^{-1}$
scheme (34)	$0.17104556531100 \times 10^{-9}$	0.99247777104929
Error of scheme (34)	$0.13329421986821 \times 10^{-9}$	$0.16280046329765 \times 10^{-1}$

Table 3 : Numerical results of Example 6

$t = 0.5$ sec	$x_1(t)$	$x_2(t)$	$x_3(t)$
Exact	0.95122942380588	$0.13887943864964 \times 10^{-10}$	$0.13887943864964 \times 10^{-10}$
RK4	0.95122939473318	$0.56965907189574 \times 10^{-10}$	$0.58351760483645 \times 10^{3}$
Error of RK4	$0.29072705243216 \times 10^{-7}$	$0.43077963324610 \times 10^{-10}$	$0.58351760483644 \times 10^{3}$
scheme (34)	0.98224764491287	$0.68582498160849 \times 10^{-5}$	$0.68582498160849 \times 10^{-5}$
Error of scheme (34)	$0.31018221106989 \times 10^{-1}$	$0.68582359281410 \times 10^{-5}$	$0.68582359281410 \times 10^{-5}$

$$\dot{x}_2 = -50x_2, \;\; x_2(0) = 1, \tag{63}$$

$$\dot{x}_3 = 70x_2 - 120x_3, \;\; x_3(0) = 2 \tag{64}$$

in the range of $0 \le t < 0.5$ sec. The exact solutions are

$$x_1(t) = \exp(-0.1t) + \exp(-50t),$$

$$x_2(t) = \exp(-50t),$$

$$x_3(t) = \exp(-50t) + \exp(-120t).$$

In the calculations we fix the stepsize to be $h = 0.025$ sec, and apply scheme (34) to the above system by letting $L = 120$ in Eq. (32), and hence $\phi = [1-\exp(-120h)]/120$. We also calculate the above system by the RK4. The calculated results were compared with the exact solutions at the final time $t = 0.5$ sec in Table 3. It can be seen that the RK4 even has a certain accurate solutions on the slowly changing components x_1 and x_2, but it induces a very large error on the rapidly changing component x_3.

5 One-dimensional cases

Specialize Eqs. (18), (22), (34) and (35) to the one-dimensional case:

$$x_{n+1} = x_n + \frac{2hx_n}{2x_n - hf_n} f_n = \frac{2x_n + hf_n}{2x_n - hf_n} x_n, \tag{65}$$

$$x_{n+1} = x_n \exp\left(\frac{hf_n}{x_n}\right), \tag{66}$$

$$x_{n+1} = x_n + \frac{2\phi x_n}{2x_n - \phi f_n} f_n = \frac{2x_n + \phi f_n}{2x_n - \phi f_n} x_n, \tag{67}$$

$$x_{n+1} = x_n \exp\left(\frac{\phi f_n}{x_n}\right). \tag{68}$$

We will apply these schemes to the one-dimensional linear model problems and give linear stability analyses.

5.1 The first linear model problem

A minimal requirement of numerical scheme for stiff differential equations is that it can pass the test of absolute stability. In this section the linear stability analyses of the above four schemes are given below. Applying the above four schemes to the first linear model problem

$$\dot{x} = \lambda x, \;\; \lambda \in \mathrm{C}^- := \{z \in \mathrm{C} |\; \mathrm{Re}(z) < 0\}, \tag{69}$$

where C is the set of complex numbers, we obtain

$$x_{n+1} = \frac{2+\lambda h}{2-\lambda h} x_n, \tag{70}$$

$$x_{n+1} = x_n \exp(\lambda h), \tag{71}$$

$$x_{n+1} = \frac{3-\exp[\mathrm{Re}(\lambda h)]}{1+\exp[\mathrm{Re}(\lambda h)]} x_n, \tag{72}$$

$$x_{n+1} = x_n \exp\left[1-\exp[\mathrm{Re}(\lambda h)]\right]. \tag{73}$$

In particular, scheme (71) is exact for Eq. (69). It is easy to verify that the corresponding four stability functions:

$$R_1(z) := \frac{2+z}{2-z}, \tag{74}$$

$$R_2(z) := \exp(z), \tag{75}$$

$$R_3(z) := \frac{3-\exp[\mathrm{Re}(z)]}{1+\exp[\mathrm{Re}(z)]}, \tag{76}$$

$$R_4(z) := \exp\left[1-\exp[\mathrm{Re}(z)]\right], \tag{77}$$

satisfy

$$|R_1(z)| \le 1, \;\; |R_2(z)| \le 1, \;\; |R_3(z)| \le 1, \;\; |R_4(z)| \le 1, \;\; \forall z \in \mathrm{C}^-. \tag{78}$$

Hence schemes (65)-(68) are all A-stable in the sense of Dahlquist (1963). Furthermore, scheme (66) is also L-stable, because its stability function satisfies $R_2(-\infty) = 0$.

5.2 The second linear model problem

Let us consider the second linear model problem:

$$\dot{x} = \lambda_0[x - p(t)] + \dot{p}(t), \;\; x(0) = x_0, \tag{79}$$

whose exact solution is $x(t) = [x_0 - p(0)]\exp(\lambda_0 t) + p(t)$. For definite we may let $p(t) = 1-\exp(\lambda_1 t)$, and assume $\lambda_0 = -10^9$ and $\lambda_1 = -1$. Hence, $x(t) = x_0\exp(-10^9 t) + [1-\exp(-t)]$, of which the first term is transient with fastly decaying, and the second term in the bracket is slowly increasing. Schemes (65) and (66) fail to calculate the solution of Eq. (79) unless the stepsize would be very small as to be explained below.

Applying schemes (67) and (68) to Eq. (79) gives good results as shown in Fig. 8(a), where the stepsize $h = 5\times10^{-4}$ sec was used, and the exact solution is compared with the numerical results, whose errors are shown in Fig. 8(b). It can be seen that the schemes (67) and (68) are not only stable but also have the accuracy in the order of $O(h)$. Instead of, the Euler method requires $h < 2/|\lambda_0| = 2\times10^{-9}$ for a stability consideration [Shampine and Gear (1979)], which in turns requires at

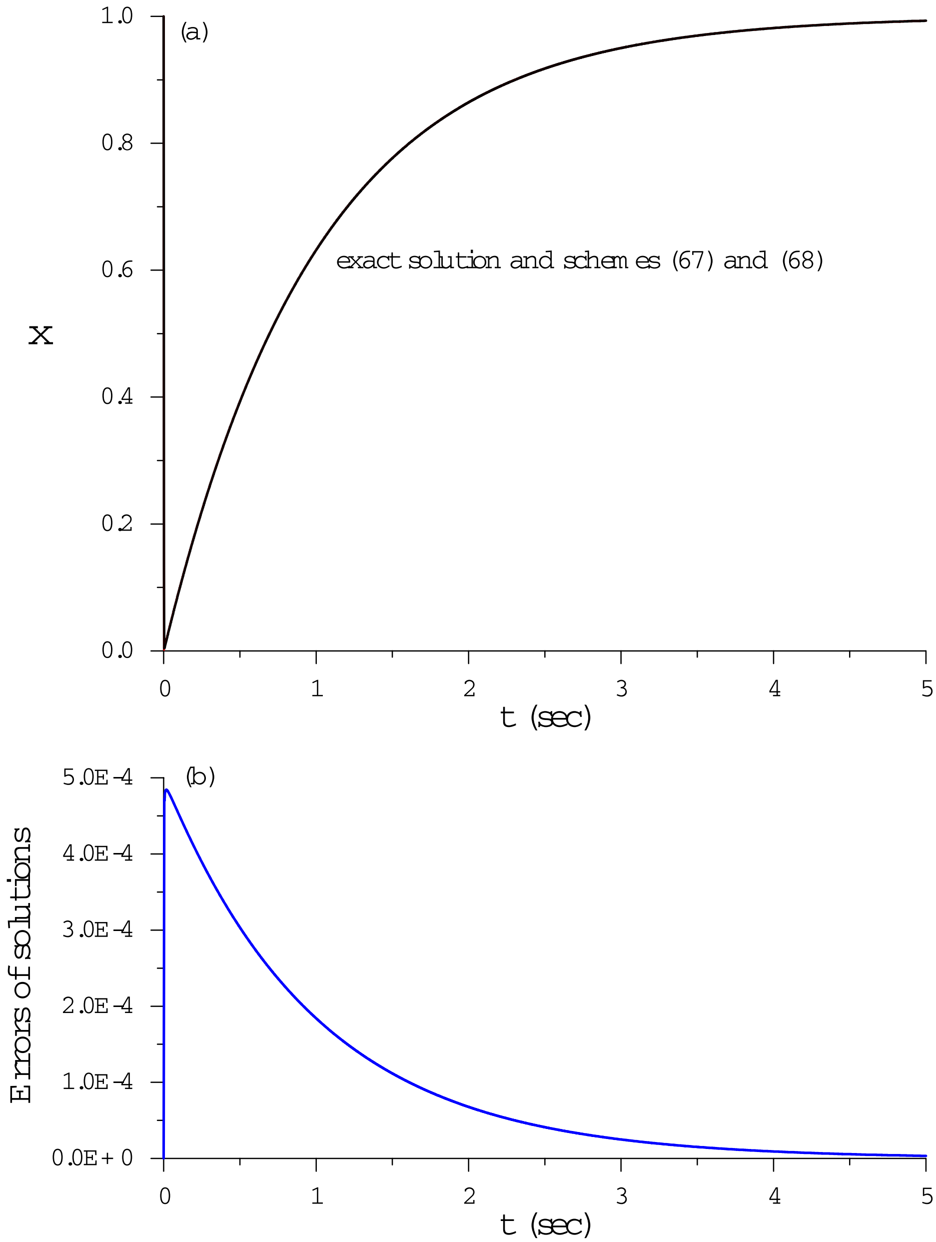

Figure 8 : The numerical solutions of the second linear model problem were obtained by schemes (67) and (68) with stepsize $h = 5 \times 10^{-4}$ sec. This model problem illustrates that some schemes which are even A-stable may fail to pass this test.

least 2.5×10^9 steps for a same calculation as that done in Fig. 8 by schemes (67) and (68).

Let us explain why scheme (67) works for this problem. In order to reply this question we shoud investigate the error propagation of scheme (67). If we define the global error as

$$\delta_n = x_n - x(t_n), \tag{80}$$

where $x(t_n)$ denotes the exact value at time $t = t_n$, from Eqs. (67) and (79) it follows that

$$\delta_{n+1} = (1 + \eta\lambda_0)\delta_n + [x(t_n) + \eta\dot{x}(t_n) - x(t_{n+1})], \tag{81}$$

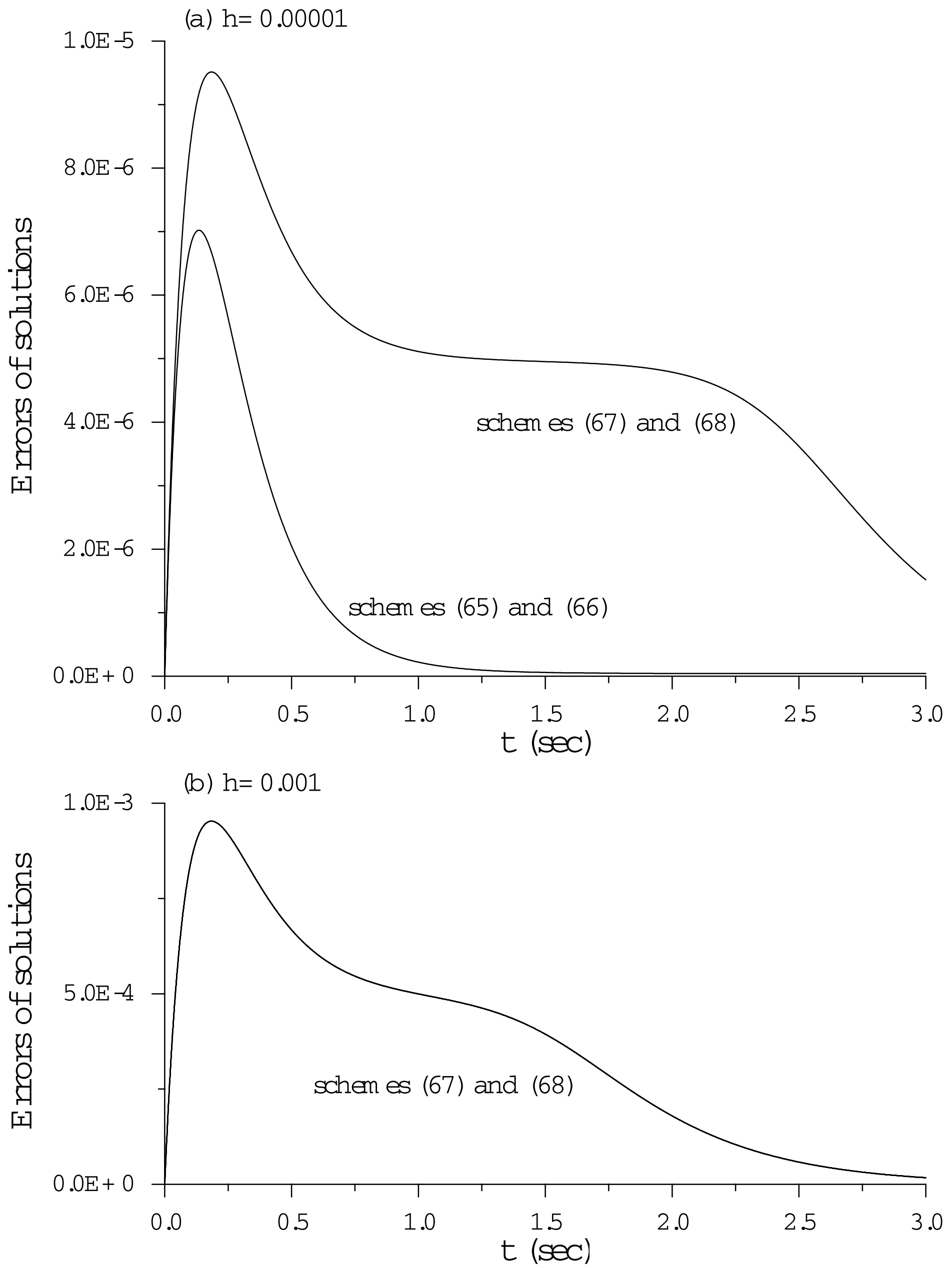

Figure 9 : The numerical solutions of Example 7 were obtained by schemes (65)-(68) with stepsize $h = 0.00001$ sec and a translation with $b = 10$. But when h increases to 0.001 sec schemes (65) and (66) fail to calculate the solutions even the same translation with $b = 10$ was considered. However, schemes (67) and (68) still work.

where

$$\eta = \frac{2\phi}{2-\phi\frac{f_n}{x_n}}, \tag{82}$$

and ϕ as defined in Eq. (32) is given by $\phi = [\exp(\lambda_0 h) - 1]/\lambda_0$. The Lipschitz constant for Eq. (79) is $-\lambda_0$; hence,

$$-\lambda_0 \geq \frac{f_n}{x_n} \geq \lambda_0. \tag{83}$$

Substituting Eq. (83) into Eq. (82) and then substituting the resultant into $1+\eta\lambda_0$ we obtain

$$1 > 1+\eta\lambda_0 > -1 \tag{84}$$

for any stepsize $h > 0$. It means that the error of scheme (67) is not amplified for any stepsize. A similar conclusion can be drawn for scheme (68) as follows. Substituting Eq. (68) for x_n into Eq. (80) we obtain the error propagation equation (81) again, but with

$$\eta = \frac{\exp\left(\frac{\phi f_n}{x_n}\right)-1}{\frac{f_n}{x_n}}. \tag{85}$$

Hence, by Eq. (83) it follows that

$$\begin{aligned} 1 &> \exp[\exp(\lambda_0 h)-1] \geq 1+\eta\lambda_0 \\ &\geq 2-\exp[1-\exp(\lambda_0 h)] > -1 \end{aligned} \tag{86}$$

for any stepsize $h > 0$, which means that the error of scheme (68) is not amplified for any stepsize.

Now, we return to scheme (65), and a similar analysis leads to

$$\eta = \frac{2h}{2-h\frac{f_n}{x_n}}. \tag{87}$$

Substituting Eq. (83) into the above, gives

$$\frac{2+3\lambda_0 h}{2+\lambda_0 h} \geq 1+\eta\lambda_0 \geq \frac{2+\lambda_0 h}{2-\lambda_0 h}. \tag{88}$$

Unless $|\lambda_0 h| < 2/3$, the error induced by scheme (65) will be amplified.

This example shows that schemes (65) and (66) can even pass the test of the first linear model problem, but they still fail to pass the test of the second linear model problem. For very stiff scalar equations the latter two schemes (67) and (68) are better than the former two schemes.

5.3 Example 7

The following stiff problem:

$$\dot{u} = 5e^{5t}(u-t)^2+1,\ \ 0 \leq t \leq 3,\ \ u(0) = -1, \tag{89}$$

has an exact solution $u(t) = t - e^{-5t}$. We first transform Eq. (89) to

$$\dot{x} = 5e^{5t}(x-b-t)^2+1,\ \ 0 \leq t \leq 3,\ \ x(0) = b-1$$

with $b = 10$. This renders $x \neq 0$ in the new equation, and then we apply schemes (65)-(68) to the above equation with $h = 0.00001$ sec. For the latter two schemes we let $L = 10$ in Eq. (32), and hence, $\phi = [1-\exp(-10h)]/10$. The errors of the above four numerical results are compared in Fig. 9(a). Then the stepsize is increased to $h = 0.001$ sec. Schemes (65) and (66) become unstable; however, schemes (67) and (68) still work, and their numerical solution errors are shown in Fig. 9(b) to be in the order of $O(h)$ and tend to decreasing as time increasing.

This example shows that the translation of state vector may give improvement of the numerical solutions for smaller stepsize; however, when a larger stepsize is used the technique of translating state vector fails to calculate the solutions.

6 Concluding remarks

At first, we have presented a numerical scheme based on the exponential mapping of the augmented form $\dot{\mathbf{X}} = \mathbf{AX}$ for nonlinear differential equations system. The group-preserving scheme (22) is better than scheme (18). However, when applied them to stiff differential equations they may fail to calculate the solutions unless the stepsize is limited to be unreasonably very small. Then, in order to tackle this difficulty of numerical instability, we have combined nonstandard finite difference method with the above group-preserving schemes and developed two nonstandard group-preserving schemes (34) and (35). They allow a larger stepsize in the calculations without inducing numerical instabilities, and also preserve the stability types of the fixed points of hyperbolic types for large stepsize as shown in Eqs. (43) and (45). The group-preserving schemes considered in this paper are all A-stable in the sense of Dahlquist. However, only schemes (34) and (35) survive for the second linear model problem, because their amplification factors of error propagation have magnitudes smaller than one for any stepsize $h > 0$.

Several numerical examples were examined, some of which were compared with exact solutions showing that the nonstandard group-preserving schemes work very well, and also have good computational efficiency and accuracy. Hence, we have unconditional stable group-preserving schemes for stiff differential equations, which provide convenient single-step explicit time integrators for stiff differential equations.

References

Brunner, H. (1974): Recursive collocation for the numerical solution of stiff ordinary differential equations. *Math. Comp.*, vol. 28, pp. 475-481.

Byrne, G. D.; Hindmarsh, A. C. (1987): Stiff ODE solvers: A review of current and coming attractions. *J. Comp. Phys.*, vol. 70, pp. 1-62.

Dahlquist, G. (1963): A special stability problem for linear multistep methods. *BIT*, vol. 3, pp. 27-43.

Gear, C. W. (1971): *Numerical Initial Value Problems in Ordinary Differential Equation*. Prentice-Hall, New Jersey.

Lapidus, L.; Schiesser, W. E. (1976): *Numerical Methods for Differential Systems*. Academic Press, New York.

Lapidus, L.; Seinfeld, J. H. (1971): *Numerical Solution of Ordinary Differential Equations*. Academic Press, New York.

Lee, H.-C.; Chen, C.-K.; Hung, C.-I. (2002): A modified group-preserving scheme for solving the initial value problems of stiff ordinary differential equations. *Appl. Math. Comp.* vol. 133, pp. 445-459.

Liu, C.-S. (2001): Cone of non-linear dynamical system and group preserving schemes. *Int. J. Non-Linear Mech.*, vol. 36, pp. 1047-1068.

Liu, C.-S.; Chang, C.-W. (2004): Lie group symmetry applied to the computation of convex plasticity constitutive equation. *CMES: Computer Modeling in Engineering & Sciences*, vol. 6, pp. 277-294.

Mickens, R. E. (1994): *Nonstandard Finite Difference Models of Differential Equations*. World Scientific, Singapore.

Mickens, R. E. (1999): Discretizations of nonlinear differential equations using explicit nonstandard methods. *J. Comp. Appl. Math.*, vol. 110, pp. 181-185.

Mickens, R. E.; Ramadhani, I. (1994): Finite-differences having the correct linear stability properties for all finite stepsizes III. *Comput. Math. Appl.*, vol. 27, pp. 77-89.

Mickens, R. E.; Smith, A. (1990): Finite-difference models of ordinary differential equations: Influence of denominator functions. *J. Franklin Institute*, vol. 327, pp. 143-149.

Robertson, H. H. (1966): The solution of a set of reaction rate equations. In Numerical Analysis, an Introduction, J. Walsh ed., pp.178-182, Academic Press, New York.

Rosenbrock, H. H.; Storey, C. (1966): *Computational Techniques for Chemical Engineers*. Pergamon Press, New York.

Shampine, L. F.; Gear, C. W. (1979): A user's view of solving stiff ordinary differential equations. *SIAM Rev.*, vol. 21, pp. 1-17.

Spijker, M. N. (1996): Stiffness in numerical initial-value problems. *J. Comp. Appl. Math.*, vol. 72, pp. 393-406.

Stabrowski, M. M. (1997): An efficient algorithm for solving stiff ordinary differential equations. *Simulation Prac. Theory*, vol. 5, pp. 333-344.