BIT Numer Math (2010) 50: 223–241
DOI 10.1007/s10543-010-0259-3

Projection methods preserving Lyapunov functions

M. Calvo · M.P. Laburta · J.I. Montijano · L. Rández

Received: 8 July 2009 / Accepted: 15 February 2010 / Published online: 5 March 2010

Abstract In this paper we consider ordinary differential equations with a known Lyapunov function. We study the use of Runge–Kutta methods provided with a dense output and a projection technique to preserve any given Lyapunov function. This approach extends previous work of Grimm and Quispel (BIT 45, 2005), allowing the use of Runge–Kutta methods for which the associated quadrature formula does not need to have positive or zero coefficients. Some numerical experiments show the good performance of the proposed technique.

Keywords Initial value problems · Lyapunov function · Numerical geometric integration · Projection methods · Explicit Runge–Kutta methods

Mathematics Subject Classification (2000) 65L05 · 65L06

Communicated by Christian Lubich.

This work was supported by D.G.I. project MTM2007-67530-C02-01.

M. Calvo · J.I. Montijano · L. Rández
IUMA, Departamento de Matemática Aplicada, Facultad de Ciencias, Universidad de Zaragoza, 50009 Zaragoza, Spain

M. Calvo
e-mail: calvo@unizar.es

J.I. Montijano
e-mail: monti@unizar.es

L. Rández
e-mail: randez@unizar.es

M.P. Laburta (✉)
IUMA, Departamento de Matemática Aplicada, Centro Politécnico Superior, Universidad de Zaragoza, 50018 Zaragoza, Spain
e-mail: laburta@unizar.es

1 Introduction

In geometric integration the development of numerical methods must be addressed not only by stability and accuracy requirements. They must also reproduce some geometric property of the differential system such as first integrals, symmetries, symplectic structure, etc. In the last years there has been great interest in this topic and a lot of work has been done about it. Information about geometric integration can be found for example in Budd & Iserles [2], Budd & Piggott [3], Hairer, Lubich & Wanner [11], Leimkuhler & Reich [15], McLachlan & Quispel [16–18] and Sanz-Serna & Calvo [23].

In this paper we will deal with autonomous initial value problems for systems of ordinary differential equations

$$y'(t) = f(y(t)), \qquad y(t_0) = y_0 \in \mathbb{R}^N, \tag{1.1}$$

having a smooth Lyapunov function $V\colon \Omega \to \mathbb{R}$, where Ω is a domain of $\mathbb{R}^N$ (see e.g. [14, 22]). More precisely, we assume that the flow of (1.1) satisfies

$$\alpha(y) := \nabla V(y) \cdot f(y) \le 0 \quad \forall y \in \Omega,$$

and the function V is bounded from below. Therefore $\dot{V} = \frac{d}{dt} V(y(t)) \le 0$ for each solution $y(t)$ of the differential system that stays in Ω. Furthermore, V is said to be a strict Lyapunov function if $<$ holds outside the set of equilibrium points $E = \{y \mid f(y) = 0\}$. We remark that we have taken autonomous flows and Lyapunov functions, but the following approach also holds for nonautonomous problems.

We are interested in numerical methods that reproduce that property of the differential system. Thus, a one-step method given by

$$y_{n+1} = \Phi_h(y_n), \quad n = 0, 1, 2, \dots$$

is said to have the Lyapunov function V if

$$V(y_{n+1}) \le V(y_n), \quad n = 0, 1, 2, \dots.$$

In this case, the Lyapunov function V is said to be strict if $<$ holds outside the set of fixed points $E_h = \{y \mid \Phi_h(y) = y\}$. Methods of low order for which a Lyapunov function V of the differential system is also a discrete Lyapunov function have been considered for example in [7, 19, 20, 26].

Grimm and Quispel in [9] proposed projection-based methods that can preserve an arbitrary smooth Lyapunov function and can have arbitrarily high order. However, this approach applied to the case of a Runge–Kutta method with coefficients (a_{ij}, b_i), has the limitation that it must have an associated quadrature formula with nonnegative coefficients $b_i \ge 0$. This implies that relevant methods, such as the well-known Dormand and Prince 5(4) pair, used in the Matlab ODE suite [25], the pair 5(4) of Fehlberg [12, II.5], or the pair 6(5) of Calvo et al. [5], cannot be suitable for this kind of problems. In this paper we present a modification of the method proposed by these authors so that, by using a dense output for the Runge–Kutta formula, any Lyapunov function can be preserved even for methods with negative coefficients b_i. Dense output for most Runge–Kutta methods of practical interest can be found in the literature (see for example [12, pp. 188], [5, 8, 21, 24]). Both the Grimm–Quispel methods as well as ours can be easily extended to systems of ordinary differential equations having more than one Lyapunov function.

The structure of this paper is as follows. In Sect. 2 we present the new approach for preserving Lyapunov functions and we analyze the order of the resulting numerical approximation. Next, in Sect. 3, we present some numerical experiments showing the performance of the proposed technique. Finally, some conclusions are established in Sect. 4.

2 Methods preserving Lyapunov functions

We start from an s-stage explicit Runge–Kutta (RK) method φ_h of order p with given coefficients a_{ij}, b_i, provided with a dense output of order $\bar{p}$, usually $\bar{p} \geq p-1$, with $\bar{s} \geq s$ stages, given by the equations:

$$y_{n+\theta} = y_n + h \sum_{i=1}^{\bar{s}} b_i(\theta) g_i, \quad \theta \in [0, 1] \tag{2.1}$$

$$g_i = f\left(y_n + h \sum_{j=1}^{i-1} a_{ij} g_j \right), \quad i = 1, \ldots, \bar{s}. \tag{2.2}$$

The proposed methods calculate an approximation y_{n+1} to the solution of (1.1) at the time $t_{n+1} = t_n + h$, from the values (t_n, y_n), according to the following three-steps algorithm:

Step 1: After calculation of the internal stages (2.2), compute $\widetilde{y}_{n+1} = \varphi_h(y_n)$, approximation given by the basic RK method:

$$\widetilde{y}_{n+1} = y_n + h \sum_{i=1}^{s} b_i(1) g_i.$$

Step 2: Compute the approximation V_{n+1} to the Lyapunov function:

$$V_{n+1} = V(y_n) + h \sum_{i=1}^{m} b_i^* \alpha(y_{n+c_i^*}),$$

where $b_i^*, c_i^*, i = 1, \ldots, m$, are respectively, the coefficients and nodes of the Gaussian quadrature formula in [0, 1] with m nodes, and the values $y_{n+c_i^*} \simeq y(t_n + c_i^* h)$, $i = 1, \ldots, m$, are calculated from (2.1)–(2.2) corresponding to the dense output of the RK method.

Step 3: Compute y_{n+1}, projection of $\widetilde{y}_{n+1}$ onto the $(N-1)$-dimensional manifold

$$M = \{y \mid g(y) := V(y) - V_{n+1} = 0\}.$$

These new methods are inspired by the projection-based methods described by Grimm and Quispel in [9]. There, the authors use in Step 2 the same RK formula as in Step 1 to approximate the Lyapunov function. In consequence, the coefficients b_i of that formula have to be non-negative, i.e. $b_i \geq 0$, $i = 1, \dots, s$. However, we do not need to impose any restriction about the coefficients of the basic RK formula. In relation with the requirements of our Step 2, it is well known that there exist Gaussian quadrature formulas with m nodes for any integer positive m, so we will easily be able to choose the most appropriate method for each situation. It will not be either difficult to get a dense output of the underlying RK method as required there.

As far as Step 3 is concerned, some known projection techniques calculate the projected value in the form

$$y_{n+1} = \widetilde{y}_{n+1} + \lambda_n w_n, \tag{2.3}$$

where $w_n = w_n(y_n, h) \in \mathbb{R}^N$ defines the direction of the projection and $\lambda_n = \lambda_n(y_n, h) \in \mathbb{R}$ is chosen so that $y_{n+1} \in M$. Thus, standard projection ([11], pp. 106) computes y_{n+1} by solving the constrained minimization problem

$$\text{Minimize } \|y_{n+1} - \widetilde{y}_{n+1}\| \quad \text{subject to } g(y_{n+1}) = 0.$$

If the norm considered is the Euclidean, $\|\cdot\|_2$, the Lagrange's necessary conditions, after approximating $\nabla g(y_{n+1}) \approx \nabla g(\widetilde{y}_{n+1})$, lead to the system

$$y_{n+1} = \widetilde{y}_{n+1} + \lambda_n \nabla g(\widetilde{y}_{n+1}) \quad \text{with } g(y_{n+1}) = 0.$$

Then we say that y_{n+1} is the orthogonal projection of $\widetilde{y}_{n+1}$ onto M, which is of the form (2.3) with $w_n = \nabla g(\widetilde{y}_{n+1})$. Another example of projection of type (2.3) can be seen in [4], where the authors propose to choose the direction w_n as the unit vector

$$w_n = \frac{\varphi_h(y_n) - \widehat{\varphi}_h(y_n)}{\|\varphi_h(y_n) - \widehat{\varphi}_h(y_n)\|_2}.$$

Here $\widehat{\varphi}_h$ denotes an embedded method to the RK method φ_h whose solution is projected onto the manifold.

After choosing the direction w_n, we have to solve the non-linear equation for λ,

$$V(\varphi_h(y_n) + \lambda w_n) - V_{n+1} = 0, \tag{2.4}$$

by using some iterative scheme, for example, by applying the Newton's method.

The new projection methods preserve the Lyapunov function V. In fact, we can write

$$V(y_{n+1}) = V_{n+1} = V(y_n) + h \sum_{i=1}^{m} b_i^* \alpha(y_{n+c_i^*}) \leq V(y_n),$$

since $h > 0$ and the coefficients $b_i^* > 0$ for all i.

If the differential system has a strict Lyapunov function V, it would be desirable that the discrete Lyapunov function corresponding to the new methods were strict too. In fact, we have

$$V(y_{n+1}) < V(y_n) \quad \Leftrightarrow \quad \sum_{i=1}^{m} b_i^* \, \alpha(y_{n+c_i^*}) < 0.$$

Since $b_i^* > 0$ for all i and $\alpha(y) \le 0$, the last inequality is violated only if all $y_{n+c_i^*}$ are critical points of (1.1), and this is quite unlikely. Note that if the solution $y(t)$ is not close to an equilibrium point, in the sense that $\|y(t) - y^*\| \ge d > 0$ for all t and for all equilibrium point y^*, then there exists $h^* = h^*(d)$ such that $\|y(t_n + c_i^* h) - y_{n+c_i^*}\| < d$ for all $h \le h^*$ and the discrete Lyapunov function is also strict.

From now on, we will denote by $y(t_n + h)$ the local solution at $t_n + h$, i.e., the solution of the differential system in (1.1) that satisfies $y(t_n) = y_n$, evaluated at $t_n + h$.

Proposition 2.1 *The local error for V calculated according to Step 2 satisfies*

$$V(y(t_n + h)) - V_{n+1} = O(h^{k+1}), \quad k = \min\{\bar{p} + 1, 2m\}.$$

Proof If we integrate $\dot{V} = \alpha(y)$ between t_n and $t_n + h$ and then we apply the Gaussian quadrature formula we obtain

$$V(y(t_n + h)) = V(y_n) + h\left[\sum_{i=1}^{m} b_i^* \alpha(y(t_n + c_i^* h)) + \frac{\psi^{2m}(\xi)}{(2m)!} \int_0^1 \prod_{i=1}^{m} (\tau - c_i^*)^2 \, d\tau\right],$$

where $\psi(\tau) = \alpha(y(t_n + \tau h))$ and $\xi \in (0, 1)$. Since we can write

$$\alpha(y(t_n + c_i^* h)) = \alpha(y_{n+c_i^*}) + O(h^{\bar{p}+1}),$$

and $\psi^{2m}(\xi) = O(h^{2m})$, the result follows. □

Now, we will give sufficient conditions that guarantee the existence of solution for λ in (2.4). They will allow us to obtain the order of the new Lyapunov function preserving projection methods. The following results follow the ideas in [4] where explicit RK methods preserving invariants were studied.

Theorem 1 *Let us consider the projection method which preserve V described above by Steps 1 to 3. If $\nabla V(y_n)^T \cdot w_n(y_n, 0) \neq 0$, then we have:*

(i) *There exists $h^* > 0$ so that (2.4) defines a unique function $\lambda_n = \lambda(y_n, h)$ for all $h \in [0, h^*]$.*
(ii) *The order of the projected method is $\ge \min\{p, \bar{p} + 1, 2m\}$.*

Proof The proof follows closely the one of Theorem 4.1, part (i), in [4], by considering the function

$$\begin{aligned} v(h, \lambda) &= V(\varphi_h(y_n) + \lambda w_n) - V_{n+1} \\ &= V(\varphi_h(y_n) + \lambda w_n) - V(y_n) - h \sum_{i=1}^{m} b_i^* \alpha(y_{n+c_i^*}), \end{aligned}$$

which is continuously differentiable in some open ball with center in $(0, 0)$ if we suppose that the functions involved are smooth. □

Remark:

1. If standard orthogonal projection is used, i.e. $w_n(y_n, h) = \nabla V(\varphi_h(y_n))$, then the sufficient condition in Theorem 1, $\nabla V(y_n)^T \cdot w_n(y_n, 0) \neq 0$ is equivalent to $\|\nabla V(y_n)\|_2^2 \neq 0$, and so it is equivalent to $\nabla V(y_n) \neq 0$.
2. If directional projection of Calvo et al. is considered [4], then

$$w_n(y_n, h) = \frac{C_{\hat{p}+1}(y_n)}{\|C_{\hat{p}+1}(y_n)\|_2} + \mathcal{O}(h),$$

where $C_{\hat{p}+1}(y_n)h^{\hat{p}+1}$ is the leading term of the local error of the RK method $\widehat{\varphi}_h$ embedded to φ_h. Hence,

$$\nabla V(y_n)^T \cdot w_n(y_n, 0) \neq 0 \Leftrightarrow \nabla V(y_n)^T C_{\hat{p}+1}(y_n) \neq 0.$$

Theorem 2 *Let us consider the projection method which preserve V described by Steps* 1 *to* 3. *If $\nabla V(\varphi_h(y_n))^T \cdot w_n(y_n, h) = B(y_n)h^r + \mathcal{O}(h^{r+1})$, with $r \geq 1$, $B(y_n) \neq 0$, and $2r \leq \min\{p, \bar{p}+1, 2m\}$, then we have*:

(i) *There exists $h^* > 0$ so that* (2.4) *defines a function $\lambda_n = \lambda(y_n, h)$ for all $h \in (0, h^*]$. Moreover, λ_n is the only function satisfying $\lambda(y_n, h) = \mathcal{O}(h^r)$.*
(ii) *The order of the projected method with λ_n is $\geq \min\{p, \bar{p}+1, 2m\} - r$.*

Proof (i) We define the real function $z(h, \mu)$ by

$$z(h, \mu) = h^{-2r}\Big[V(\varphi_h(y_n) + \mu h^r w_n) - V_{n+1}\Big], \quad \text{for } h \neq 0.$$

By applying Taylor's theorem and taking into account the hypothesis imposed, we can write

$$\begin{aligned} z(h, \mu) &= h^{-2r}\Big[V(\varphi_h(y_n)) + \mu h^r \nabla V(\varphi_h(y_n))^T w_n \\ &\qquad + \frac{\mu^2 h^{2r}}{2} w_n^T H_V(\varphi_h(y_n)) w_n + \mathcal{O}(h^{3r}) - V_{n+1}\Big] \\ &= \mu B(y_n) + \frac{\mu^2}{2} w_n^T H_V(\varphi_h(y_n)) w_n + \mathcal{O}(h) \end{aligned}$$

where H_V is the Hessian matrix. Thus, we can define

$$z(0, \mu) := \lim_{h \to 0} z(h, \mu) = \mu B(y_n) + \frac{\mu^2}{2} w_n(y_n, 0)^T H_V(y_n) w_n(y_n, 0)$$

and $z(h, \mu)$ is continuous for $h \geq 0$. Furthermore, the function $\partial z / \partial \mu$ is also continuous in a neighborhood of $(0, 0)$, and we have

$$z(0, 0) = 0, \qquad \frac{\partial z}{\partial \mu}(0, 0) = B(y_n) \neq 0.$$

Therefore, the implicit function theorem guarantees the existence of $h^* > 0$ and a unique function $\mu = \mu(h)$ such that $\mu(0) = 0$ and $z(h, \mu(h)) = 0$ for all $h \in [0, h^*]$. This means that $\lambda_n(h) = h^r \mu(h)$ is the only solution of (2.4) for all $h \in (0, h^*]$ satisfying $\lambda_n(h) = \mathcal{O}(h^r)$.

(ii) To know the order of the projected method, we write

$$z(h, \mu) = z(h, 0) + \frac{\partial z}{\partial \mu}(h, 0)\mu + \mathcal{O}(\mu^2),$$

where

$$z(h, 0) = h^{-2r}\Big[V(\varphi_h(y_n)) - V_{n+1}\Big] = \mathcal{O}(h^{l-2r+1}),$$

$$\frac{\partial z}{\partial \mu}(h, 0) = \frac{\partial z}{\partial \mu}(0, 0) + \mathcal{O}(h) = B(y_n) + \mathcal{O}(h)$$

with $l = \min\{p, \bar{p} + 1, 2m\}$. It implies that $\mu = \mathcal{O}(h^{l-2r+1})$ and so $\lambda_n = \mathcal{O}(h^{l-r+1})$. Therefore, the order of the projected method is $\geq l - r$. □

Remark The above theorem guarantees the existence of a solution $\lambda(y_n, h)$ for h small enough, but uniqueness cannot be ensured in general if $\nabla V(y_n)^T \cdot w_n(y_n, 0) = 0$.

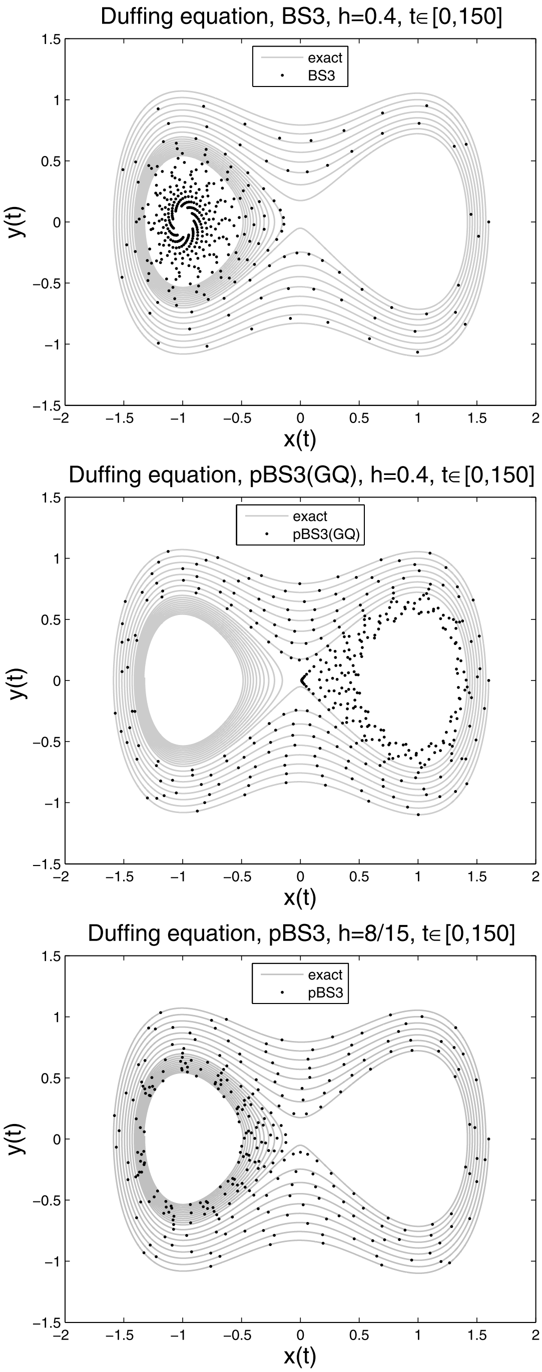
3 Numerical experiments

We now present some numerical experiments to show the behavior of the new projection methods which preserve Lyapunov functions.

The V-preserving methods considered here are based on well known explicit Runge–Kutta formulas, used in the Matlab package [25]. They have been implemented by using either fixed or variable step size strategies. Thus,

- BS32 will denote the 4-stage embedded RK pair of order 3(2) derived by Bogacki and Shampine [1]. The step size control is done taking into account the estimate of the local error, obtained by the difference between the 3th-order and the 2nd-order approximations [12, II.4].
- pBS32 denotes the method preserving V obtained from BS32 as described in Sect. 2. It uses a dense output of order 3 with no additional f-function evaluations over the pair, given in $[t_n, t_{n+1}]$ by the Hermite interpolant obtained from the values $(t_n, y_n, f(y_n))$ and $(t_{n+1}, y_{n+1}, f(y_{n+1}))$. The method pBS32 also uses the Gaussian quadrature formula in [0, 1] with 2 nodes.
- BS3 represents the 3-stage 3th order method in the pair BS32, implemented with fixed step size.
- pBS3 is used to represent the method preserving Lyapunov functions obtained as described in the previous item, but now taking BS3 as the basic method instead of BS32.
- DP54 will denote the 7-stage embedded 5(4) RK pair of Dormand and Prince [6], of course implemented with variable step size.

Fig. 1 Duffing equation (3.1): phase portraits, $t \in [0, 150]$



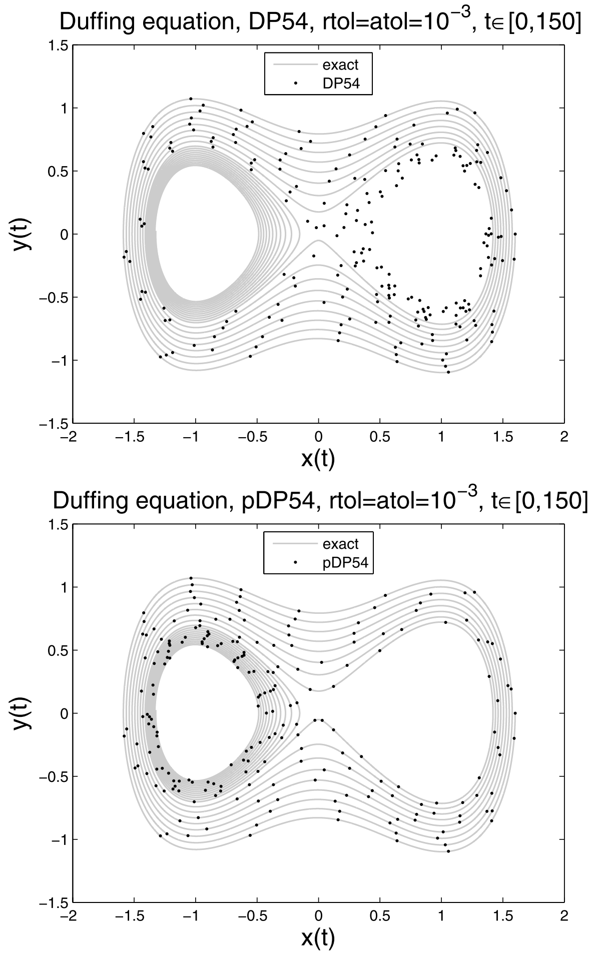


Fig. 2 Duffing equation (3.1): phase portraits, $rtol = atol = 10^{-3}$

- pDP54 represents the method that preserves Lyapunov functions, obtained from DP54. In this case, it is possible to construct a continuous formula of order 4 that does not require any additional function evaluation of the form (2.1) with $\bar{s} = 6$, where g_i, $i = 1, \ldots, 6$, are the function evaluations corresponding to the 5th order method. For this projected method pDP54 we have used the Gaussian quadrature formula with 3 nodes.
- DP5 is the 6-stage formula in DP54 of order 5, implemented with a fixed step size.
- pDP5 is used to represent the V-preserving method obtained as pDP54, but now from the underlying formula DP5.

The numerical results presented here correspond to standard orthogonal projection. Thus, we have taken the unit vector $w_n = \nabla V(\tilde{y}_{n+1})/\|\nabla V(\tilde{y}_{n+1})\|_2$ as the direction in the step from (t_n, y_n).

Let us notice that the coefficients b_i of the formulas in BS32 are all positive. Therefore, the projected method proposed by Grimm and Quispel in [9] can be considered

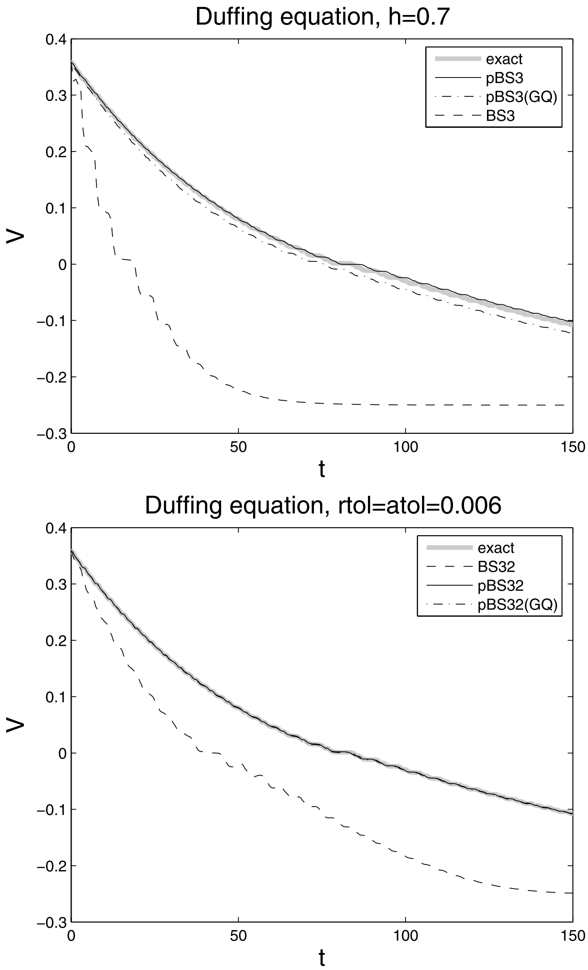


Fig. 3 Duffing equation (3.1): $V(y(t))$ against t, $h=0.7$ (*top*), $rtol=atol=0.006$ (*bottom*)

for that pair. We will refer to that method as pBS32(GQ) or pBS3(GQ), respectively, according to whether variable or fixed step size is used.

Note also that pBS3 uses 4 function evaluations per step because to compute the dense output, the FSAL stage $f(\widetilde{y}_{n+1})$ must be calculated, whereas BS3 and pBS3(GQ) only require 3 function evaluations per step.

The first test problem is the Duffing equation without external forcing given by [10]:

$$x'' + ax' - x + x^3 = 0, \tag{3.1}$$

with a constant $a \geq 0$, which can be written as the first-order system

$$x' = y, \qquad y' = x - x^3 - ay. \tag{3.2}$$

Since $V(x,y) = y^2 - x^2 + \frac{1}{2}x^4$ has $\dot{V} = -2ay^2 \leq 0$, it is a Lyapunov function for (3.2).

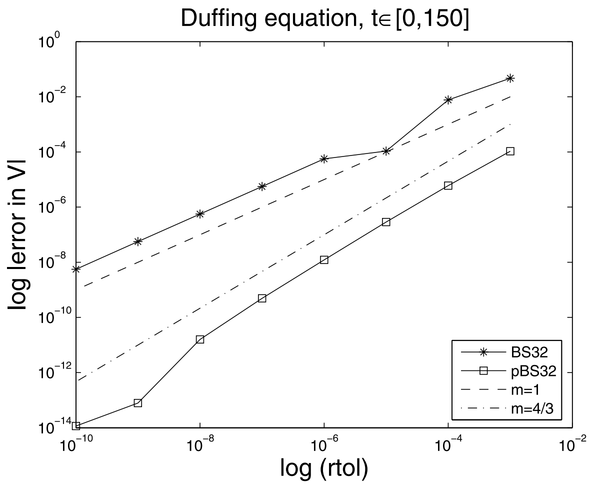


Fig. 4 Duffing equation (3.1): error in V against *rtol*, $m = 1$ and $m = 4/3$ are reference straight lines with slopes 1 and 4/3, respectively

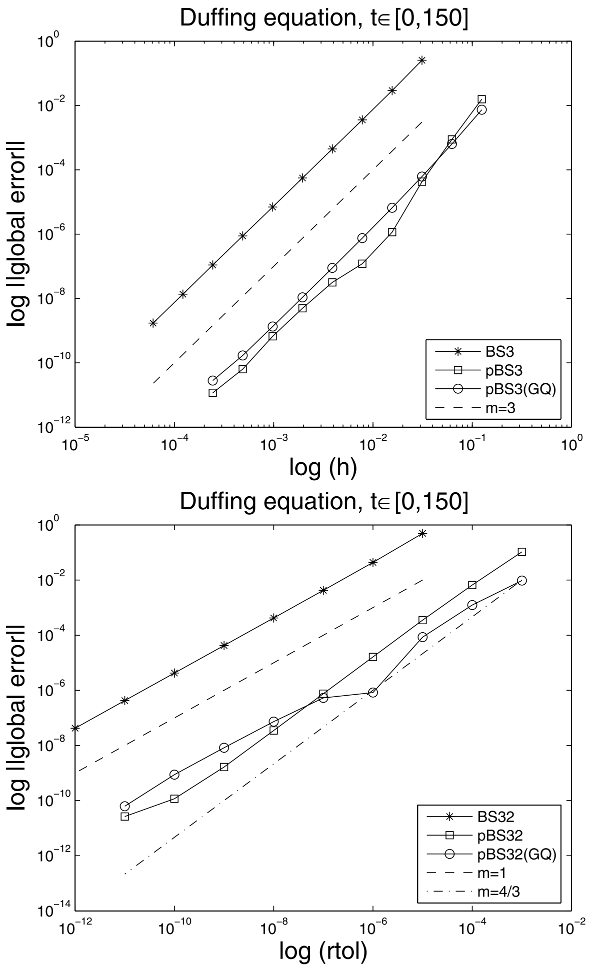


Fig. 5 Duffing equation (3.1): global error against h (*top*) or *rtol* (*bottom*), $m = 3$, $m = 1$ and $m = 4/3$ are reference straight lines with slopes 3, 1 and 4/3, respectively

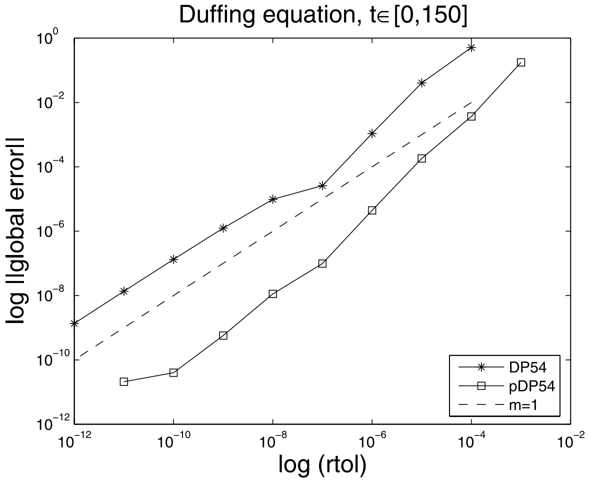


Fig. 6 Duffing equation (3.1): global error against *rtol*, $m = 1$ is a reference straight line with slope 1

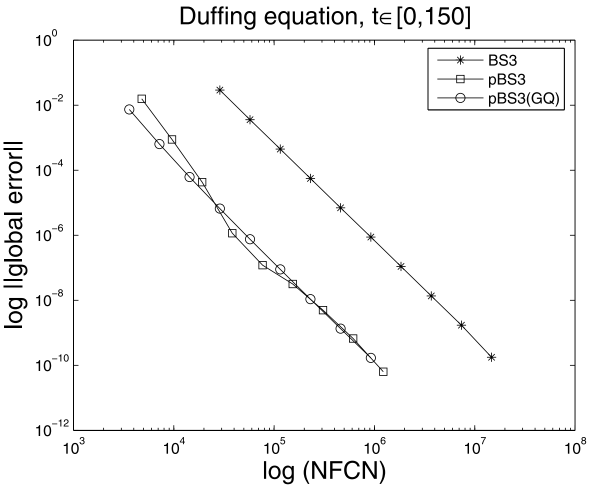
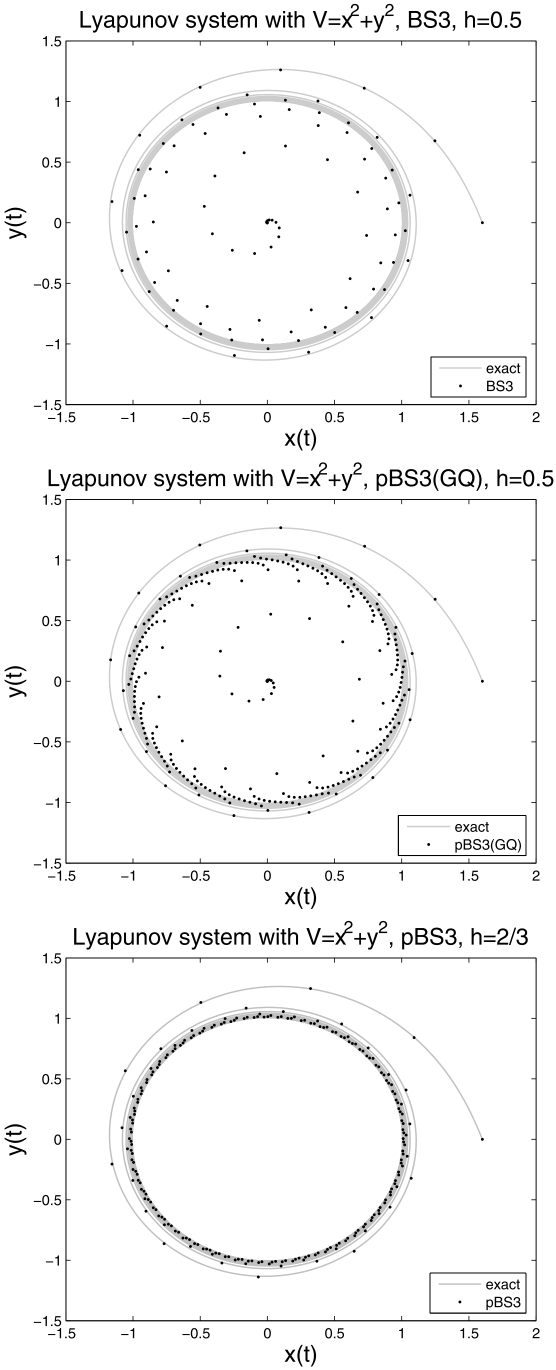


Fig. 7 Duffing equation (3.1): global error against function evaluations

In the numerical experiments we have taken as initial condition at $t = 0$ the point $y_0 = (1.6, 0)$, and as value of the parameter $a = 0.01$. The integrations have been carried out for $t \in [0, 150]$.

We show in Fig. 1 the phase portraits corresponding, respectively, to BS3, pBS3(GQ) and pBS3, each one compared to the exact solution. We have taken the step size $h = 0.4$ for BS3 and pBS3(GQ), and $h = 8/15$ for pBS3, to compare methods at equal work. Recall that pBS3 employs 4 function evaluations per step whereas the other two methods only employ 3. The phase portrait is incorrect as for BS3, whose numerical solution goes to the equilibrium point $(-1, 0)$ in a wrong way, as even for pBS3(GQ), which ends up at the wrong equilibrium point $(1, 0)$. However, the new projection method pBS3 already exhibits the proper behavior even though it uses a larger step size. It can also be seen that if we reduce the step size h and we take $h = 0.3$, then the projection method pBS3(GQ) also gives rise to a correct phase

Fig. 8 Second test problem (3.3): phase portraits, $t \in [0, 150]$



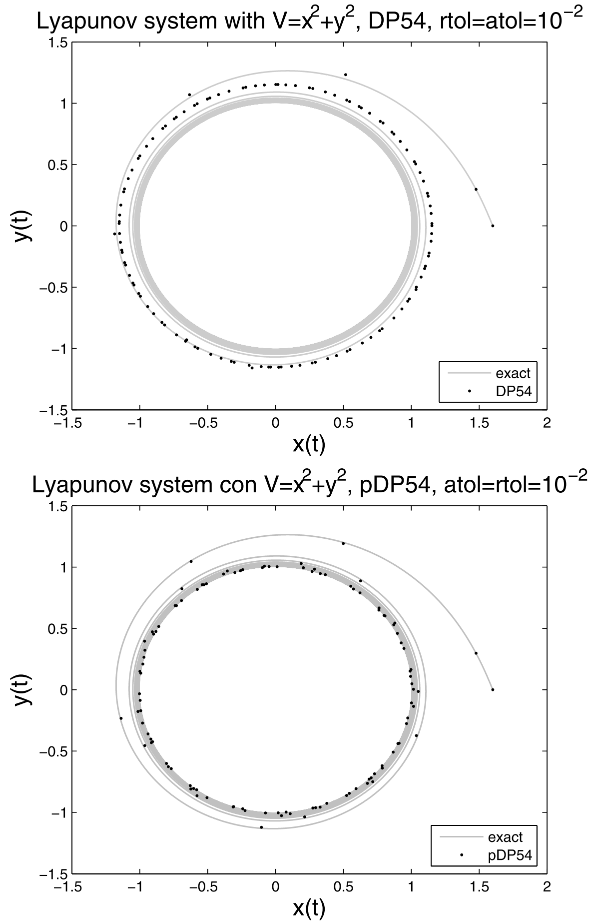


Fig. 9 Second test problem (3.3): phase portraits, $rtol = atol = 10^{-2}$, $t \in [0, 150]$

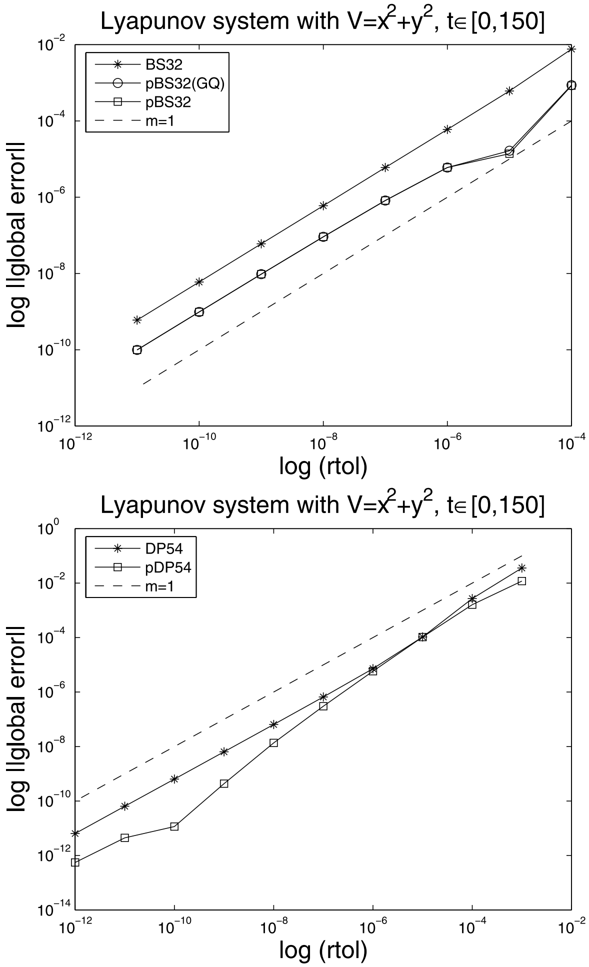
portrait, whereas the standard method BS3 still shows an incorrect behavior going towards the point $(1, 0)$.

This good behavior of the new projection methods has been also observed for variable stepsize. In Fig. 2 the results for DP54 and pDP54 can be seen, both for relative and absolute error tolerances $rtol = atol = 10^{-3}$.

In Fig. 3 we represent the approximation of the Lyapunov function V against time t. On the left-hand side, the new method pBS3 is already exhibiting the correct behavior for the large step size $h = 0.7$. It is closely followed by the other V-preserving method pBS3(GQ), whereas the faster decay corresponds to the standard method BS3. On the right-hand side, for variable step size with $rtol = atol = 0.006$, the two projected methods, pBS32 and pBS32(GQ), have a similar behavior, both closed to the exact solution, and it is much better than the corresponding to the standard pair BS32.

In Fig. 4 the error in V at the final point of the integration interval against $rtol$ is represented for BS32 and pBS32 in a log-log scale. There, we can see again the superior performance of the projection method. The slope $4/3$ of the auxiliary straight

Fig. 10 Second test problem (3.3): global error against *rtol*, $m = 1$ is a reference straight line with slope 1



line indicates that this method behaves as if it had order 4, according to the tolerance proportionality theory [13].

The two graphics in Fig. 5 show the performance of the global error at the end of the integration against either the step size h or *rtol*, respectively, in a log-log scale. Once again we observe the superior performance of the projection methods over the standard ones. The slopes of the reference straight lines show that all the methods have order 3, except our pBS32 that behaves as a fourth-order method for variable step size. Analogous results have been obtained for DP54 and its projected method pDP54, as it can be appreciated in Fig. 6. For small tolerances, both methods behave like 5th-order methods as expected.

Figure 5 shows that projection methods give smaller errors than the standard, non-projected, method, and the new projection technique gives slightly smaller errors that those of Grimm and Quispel. To compare the efficiency of the methods with the same work, we display in Fig. 7 the global error against the number of function evaluations

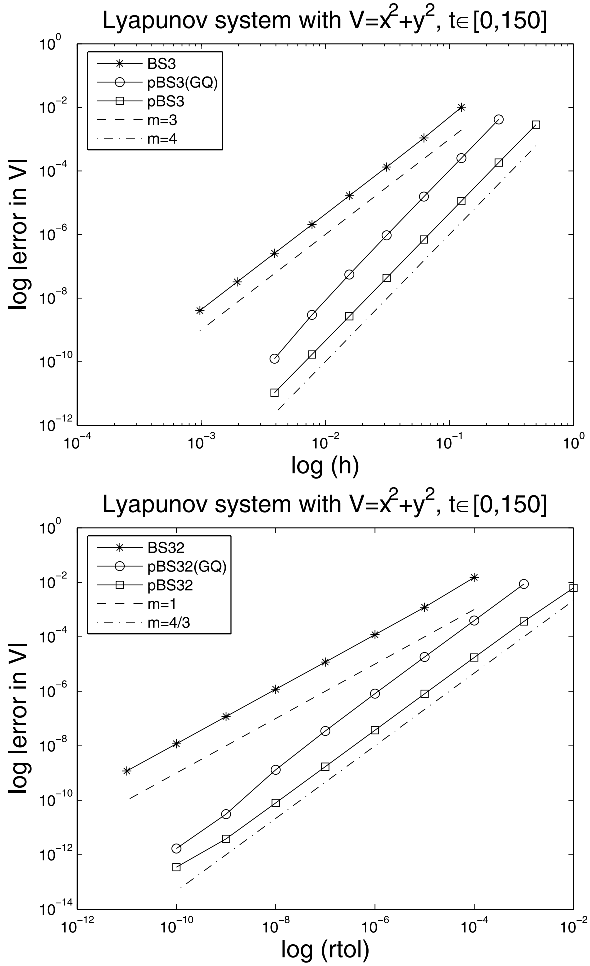


Fig. 11 Second test problem (3.3): error in V against h (*top*) or *rtol* (*bottom*), $m=3$, $m=4$, $m=1$ and $m=4/3$ are reference straight lines with slopes 3, 4, 1 and 4/3, respectively

for BS3, pBS3 and pBS3(GQ) when $t \in [0, 150]$. Now, the efficiency of the two projection methods is very similar, and they perform more efficiently than the standard method BS3.

Our second test problem is given by the equations:

$$x' = -y - x(1 - \sqrt{x^2 + y^2})^2,$$
$$y' = x - y(1 - \sqrt{x^2 + y^2})^2. \tag{3.3}$$

This system has the Lyapunov function $V(x, y) = x^2 + y^2$. We have taken $y_0 = (1.6, 0)$ as initial condition, and [0, 150] as the integration interval. As is pointed out in [9], the exact solution starting outside the circle centered in $(0, 0)$ with radius one is attracted to that circle, staying there later on.

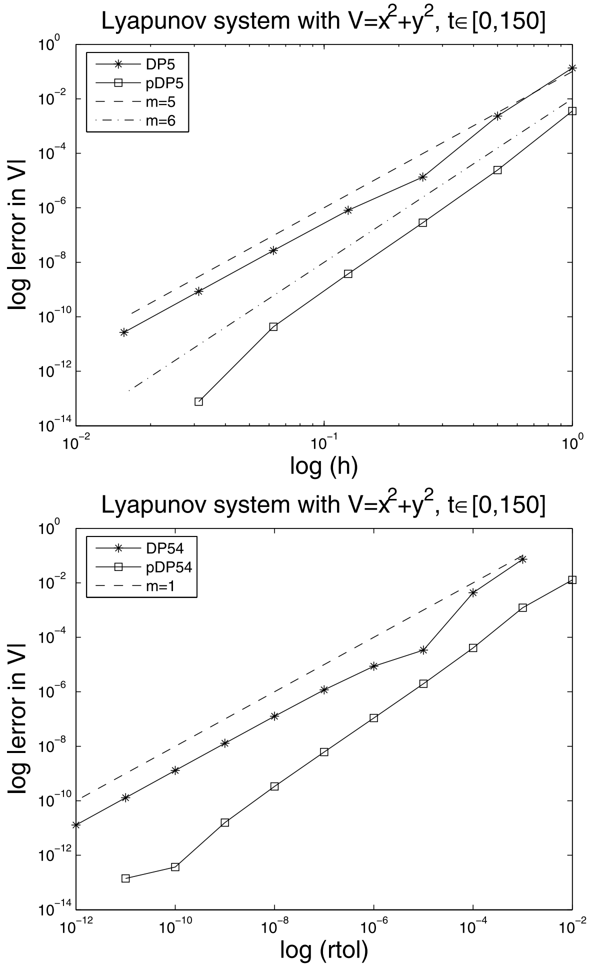


Fig. 12 Second test problem (3.3): error in V against h (*top*) or *rtol* (*bottom*), $m=5$, $m=6$ and $m=1$ are reference straight lines with slopes 5, 6 and 1, respectively

In Fig. 8 we show the phase portraits obtained, respectively, for BS3, pBS3(GQ) and pBS3. In order to compare them at equal work, we have taken the step size $h=0.5$ for the two first methods, and $h=2/3$ for the third one. There we can see that only pBS3 exhibits the right behavior. If smaller step sizes are used, the other two methods also provide the right phase portrait, but the standard method BS3 always requires smaller step sizes than the projection methods to get it. Figure 9 shows the results obtained for variable step size with error tolerances $rtol = atol = 10^{-2}$ for DP54 and pDP54. They are wrong for the standard method and correct for the projected method preserving V.

In Fig. 10 we have represented the Euclidean norm of the global error at the end of the integration versus the error tolerance, respectively for BS32 and DP54, each one together with its projected methods. In the left graphic window, the two V-preserving methods obtain practically the same results, and they are better than those of BS32. The auxiliary straight line with slope one shows that the global error behaves in the three methods as $\mathcal{O}(tol)$. In the right graphic window we observe the superior performance of pDP54 over DP54, more noticeable for small tolerances.

Graphics in Fig. 11 represent the error in the Lyapunov function at the end point versus either the stepsize h or the error tolerance $rtol$ for the Bogacki–Shampine method and the V-preserving methods constructed from it. In both cases, our method gives the better results. It can be seen that the projection methods behave as 4th-order methods in the V-error. Analogous comments can be done about the numerical results in Fig. 12 for the Dormand and Prince method and the methods preserving V derived from it.

4 Conclusions

In this paper new projected methods preserving Lyapunov functions have been studied. They are based on an explicit RK method provided with dense output, together with a quadrature formula with positive weights. The preservation of the Lyapunov function is achieved after projecting at each step the numerical solution onto the appropriate manifold. These new methods are a modification of the Grimm and Quispel ones in [9], but they do not require the non-negativity of the coefficients b_i of the underlying RK method. Numerical experiments have highlighted the superior performance of the new methods over the standard ones.

Acknowledgements The authors thank the referees for the careful reading of the manuscript and their helpful suggestions.

References

1. Bogacki, P., Shampine, L.F.: A 3(2) pair of Runge–Kutta formulas. Appl. Math. Lett. **2**(4), 321–325 (1989)
2. Budd, C.J., Iserles, A. (eds.): Geometric integration: numerical solution of differential equations on manifolds. Special edition of Philos. Trans. R. Soc. Lond. **357** (1999)
3. Budd, C.J., Piggott, M.D.: Geometric integration and its applications. In: Proceedings of ECMWF Workshop on "Developments in Numerical Methods for Very High Resolution Global Models", pp. 93–118 (2000)
4. Calvo, M., Hernández-Abreu, D., Montijano, J.I., Rández, L.: On the preservation of invariants by explicit Runge–Kutta methods. SIAM J. Sci. Comput. **28**(3), 868–885 (2006)
5. Calvo, M., Montijano, J.I., Rández, L.: A new embedded pair of Runge–Kutta formulas of order 5 and 6. Comput. Math. Appl. **20**(1), 15–24 (1990)
6. Dormand, J.R., Prince, P.J.: A family of embedded Runge–Kutta formulae. J. Comput. Appl. Math. **6**(1), 19–26 (1980)
7. Elliot, C.M., Stuart, A.M.: The global dynamics of discrete semilinear parabolic equations. SIAM J. Numer. Anal. **30**, 1622–1663 (1993)
8. Enright, W.H., Jackson, K.R., Nørsett, S.P., Thomsen, P.G.: Interpolants for Runge–Kutta formulas. ACM Trans. Math. Softw. **12**(3), 193–218 (1986)
9. Grimm, V., Quispel, G.R.W.: Geometric integration methods that preserve Lyapunov functions. BIT Numer. Math. **45**, 709–723 (2005)
10. Guckenheimer, J., Holmes, P.: Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields. Springer, New York (1983)
11. Hairer, E., Lubich, C., Wanner, G.: Geometric Numerical Integration: Structure Preserving Algorithms for Ordinary Differential Equations. Springer, Berlin (2002)

12. Hairer, E., Nørsett, S.P., Wanner, G.: Solving Ordinary Differential Equations I, Nonstiff Problems, 2nd edn. Springer, Berlin (1993)
13. Higham, D.J.: The tolerance proportionality of adaptive ODE solvers. J. Comput. Appl. Math. **45**, 227–236 (1993)
14. Lakshmikantham, V., Matrosov, V.M., Sivasundaram, S.: Vector Lyapunov Functions and Stability Analysis of Nonlinear Systems. Kluwer, Dordrecht (1991)
15. Leimkuhler, B., Reich, S.: Simulating Hamiltonian Dynamics. Cambridge University Press, Cambridge (2004)
16. McLachlan, R.I., Quispel, G.R.W.: Six lectures on the geometric integration of ODEs. In: DeVore, R.A., et al. (eds.) Foundations of Computational Mathematics, pp. 155–210. Cambridge University Press, Cambridge (2001)
17. McLachlan, R.I., Quispel, G.R.W.: Splitting methods. Acta Numer. **11**, 341–434 (2002)
18. McLachlan, R.I., Quispel, G.R.W.: Geometric integrators for ODEs. J. Phys. A **39**(19), 5251–5285 (2006)
19. McLachlan, R.I., Quispel, G.R.W., Robidoux, N.: A unified approach to Hamiltonian systems, Poisson systems, gradient systems and systems with a Lyapunov function and/or first integrals. Phys. Rev. Lett. **81**, 2399–2403 (1998)
20. McLachlan, R.I., Quispel, G.R.W., Robidoux, N.: Geometric integration using discrete gradients. Philos. Trans. R. Soc. Lond. A **357**, 1021–1045 (1999)
21. Owren, B., Zennaro, M.: Derivation of efficient, continuous, explicit Runge–Kutta methods. SIAM J. Sci. Comput. **13**, 1488–1501 (1992)
22. Rouche, N., Habets, P., Laloy, M.: Stability Theory by Liapunov's Direct Method. Applied Mathematical Sciences, vol. 22. Springer, New York (1977)
23. Sanz-Serna, J.M., Calvo, M.P.: Numerical Hamiltonian Problems, Applied Mathematics and Mathematical Computation, vol. 7. Chapman & Hall, London (1994)
24. Shampine, L.F.: Interpolation for Runge–Kutta formulas. SIAM J. Numer. Anal. **22**, 1014–1027 (1985)
25. Shampine, L.F., Reichelt, M.W.: The MATLAB ODE suite. SIAM J. Sci. Comput. **18**(1), 1–22 (1997)
26. Stuart, A.M., Humphries, A.R.: Dynamical Systems and Numerical Analysis. Cambridge University Press, Cambridge (1996)