Some Remarks on Numerical Methods for Second Order Differential Equations on the Orthogonal Matrix Group

Nicoletta Del Buono, Cinzia Elia

Dipartimento Interuniversitario di Matematica,
Università degli Studi di Bari
Via E.Orabona, 4 - I70125 Bari ITALY
[delbuono,elia]@dm.uniba.it

Abstract. In the last years several numerical methods have been developed to integrate matrix differential equations which preserve certain features of the theoretical solution such as orthogonality, eigenvalues, first integrals, etc. In this paper we approach the numerical solution of a second order matrix differential system whose solution evolves on the Lie- group of the orthogonal matrices $\mathcal{O}_n$. We study the orthogonality properties of classical Runge Kutta Nyström methods and non standard numerical procedures for second order ordinary differential equations.

1 Introduction

Recently, there has been an increasing interest in conservative numerical methods for solving ordinary differential equations which preserve certain features of the theoretical solution such as orthogonality, symplecticness, isospectrality, first integrals, etc. ([1], [2], [3], [10]). In this paper we shall concern with a system of special second order ordinary differential equations (ODEs) of dimension n, whose solutions remain for all t on the Lie group of the orthogonal matrices

$$\mathcal{O}_n = \{Y \in \mathbb{R}^{n\times n} |\ Y^T Y = I_n\},$$

where I_n is the unity matrix of dimension n.

Particularly, we are interested in solving differential equations of the following form:

$$\ddot{Y}(t) = C(t, Y(t))Y(t), \qquad t > 0, \qquad Y(0) = I_n, \qquad \dot{Y}(0) = B, \qquad (1)$$

where B is a skew-symmetric matrix (i.e. $B^T = -B$) and the matrix function $C : \mathbb{R} \times \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ is such that $Y(t) \in \mathcal{O}_n$, for all $t > 0$.

Second order ordinary differential equations evolving on $\mathcal{O}_n$ arise in several applications, for example in computation of embedded geodesics curve (see [5]). If we set $P(t) = \dot{Y}(t)$ the differential system (1) can be transformed into the equivalent first order system of twice the dimension

$$\begin{pmatrix} \dot{P}(t) \\ \dot{Y}(t) \end{pmatrix} = \begin{pmatrix} 0 & C(t, Y(t)) \\ I & 0 \end{pmatrix} \begin{pmatrix} P(t) \\ Y(t) \end{pmatrix}, \qquad \begin{pmatrix} P(0) \\ Y(0) \end{pmatrix} = \begin{pmatrix} B \\ I \end{pmatrix}, \qquad (2)$$

P.M.A. Sloot et al. (Eds.): ICCS 2002, LNCS 2331, pp. 467–475, 2002.

thus usual methods for first order ODEs may be applied. However orthogonal preserving methods applied to (2) do not preserve the orthogonality of $Y(t)$. Moreover, it is well known that a direct solution of (2) may give computational advantages and standard numerical methods for second order ODEs are often considered to be more efficient for systems of the form (1) (see [6], [9]). In this paper after recalling some concepts on the geometrical structure of the Lie group $\mathcal{O}_n$, we particularize the second order differential equation we are interested in. In Section 3 we study when a general s-stage Runge Kutta Nyström method (hereafter abbreviated as RKN) is an orthogonal integrator for equation (1). In Section 4, we discuss the Cayley approach applied to second order orthogonal equations. Finally, we present some numerical tests to illustrate the behavior of the algorithms.

2 Background

In this section we review some concepts from differential geometry and provide the geometrical structure of the group of orthogonal matrices, which will be used throughout the rest of the paper. We also characterize a kind of second order ODEs on $\mathcal{O}_n$.

To begin with, assume $Y(t) \in \mathcal{O}_n$ for all t and denote by $\mathcal{T}_{Y(t)}\mathcal{O}_n$ the tangent space at $Y(t)$. The equation defining a tangent vector to $\mathcal{O}_n$ at the point Y is easily obtained by differentiating the constraint $YY^T = I$, i.e.

$$\dot{Y}Y^T + Y\dot{Y}^T = 0,$$

hence:

$$\mathcal{T}_Y\mathcal{O}_n = \{\Delta \in \mathbb{R}^{n\times n} | \Delta Y^T + Y\Delta^T = 0\},$$

Clearly, the Lie algebra o_n of $\mathcal{O}_n$ (i.e. the tangent space at the identity) is the set of all skew symmetric matrices:

$$o_n = \{B \in \mathbb{R}^{n\times n} | B^T + B = 0\}.$$

Furthermore, by differentiating twice the constraint, we get

$$\ddot{Y}Y^T + 2\dot{Y}\dot{Y}^T + Y\ddot{Y}^T = 0.$$

that is $\ddot{Y}$ belongs to the set:

$$\mathcal{N}_Y\mathcal{O}_n = \{\Omega \in \mathbb{R}^{n\times n} | \Omega Y^T + Y\Omega^T + 2\Delta\Delta^T = 0, \text{ with } \Delta \in \mathcal{T}_Y\mathcal{O}_n\}.$$

Theorem 1. *Let $Y(t)$ be the solution of (1). Then $Y(t)$ belongs to $\mathcal{O}_n$ for all $t > 0$, if and only if there exists $A : \mathbb{R} \times \mathbb{R}^{n\times n} \to o_n$ continuous and locally Lipschitz skew-symmetric matrix function such that*

$$C(t, Y) = \dot{A}(t, Y) + A_Y(t, \dot{Y}) + A(t, Y)A(t, Y), \tag{3}$$

with $A(0, Y(0)) = B$, where $\dot{A}$ denotes the derivative with respect to t and $A_Y(t, X) = \sum_{i,j=1}^{n} \frac{\partial A(t, Y)}{\partial Y_{i,j}} X_{i,j}$.

Proof. Suppose that $C(t,Y)$ in (1) is given by (3), then being $A(t,V)$ a skew-symmetric matrix function, the first order differential system

$$\dot{V}(t) = A(t, V(t))V(t), \qquad V(0) = I_n, \tag{4}$$

has a solution $V(t)$ which is an orthogonal matrix for all $t > 0$. Moreover, by differentiating (4), we obtain that $V(t)$ satisfies the second order differential system:

$$\begin{array}{l} \ddot{V}(t) = [\dot{A}(t,V) + A_V(t,\dot{V}) + A(t,V)A(t,V)]V(t), \\ V(0) = I_n, \qquad \dot{V} = B. \end{array} \tag{5}$$

From the uniqueness of the solution of (1) and (4), it follows that $V(t) = Y(t)$, for all $t > 0$. Now we assume that the solution of (1) is orthogonal. Then it satisfies a ordinary differential equation of the form

$$\dot{Y}(t) = A(t, Y(t))Y(t), \qquad Y(0) = I_n, \tag{6}$$

where $A : \mathbb{R} \times \mathbb{R}^{n\times n} \to o_n$ is continuous and locally Lipschitz skew-symmetric matrix function. Then the results follows by differentiating (6).

Now the following result may be easily proved.

Lemma 1. *Suppose that the unique solution $Y(t)$ of (1) is orthogonal and the matrix function $C(t)$ is independent on Y. Then $C(t) = \dot{A}(t) + A(t)A(t)$ for all $t > 0$, with $A(t)$ the skew-symmetric matrix function given by*

$$A(t) = B + \frac{1}{2}\int_0^t [C(s) - C^T(s)]\mathrm{ds}. \tag{7}$$

Proof. From the Theorem 1 it follows that there exists a skew-symmetric matrix function $A(t)$ such that:

$$C(t) = \dot{A}(t) + A(t)A(t), \quad t > 0. \tag{8}$$

Hence

$$-C^T(t) = -\dot{A}^T(t) - A^T(t)A^T(t). \tag{9}$$

Adding (8) to (9) and using the skew-symmetry of A and $\dot{A}$, then (7) follows.

Remark 1. From Theorem 1 it follows that the orthogonal solution of the second order differential equation (1) with $C(t)$ depending only on t, is equivalent to the solution of the first order differential equation of the same dimension. This leads to a computational advantage.

3 RKN methods and orthogonality

Let $h > 0$ be the step-size, $\{t_k\}$ the set of the step points and $Y_{k+1}, \dot{Y}_{k+1}$ denote the numerical approximation of $Y(t_{k+1})$ and $\dot{Y}(t_{k+1})$, respectively. A s-stage

RKN method for (1) is given by:

$$\begin{aligned} Y_{k+1} &= Y_k + h\dot{Y}_k + h^2 \sum_{i=1}^{s} \bar{b}_i K_i', \\ \dot{Y}_{k+1} &= \dot{Y}_k + h \sum_{i=1}^{s} b_i K_i', \end{aligned} \tag{10}$$

with

$$K_i' = C(t_k + c_i h, Y_k + c_i h\dot{Y}_k + h^2 \sum_{j=1}^{s} \bar{a}_{ij} K_j')(Y_k + c_i h\dot{Y}_k + h^2 \sum_{j=1}^{s} \bar{a}_{ij} K_j'),\ i = 1\ldots, s$$

where a_{ij}, b_i, c_i, for $i, j = 1, \ldots, s$ are real coefficients. Furthermore, introducing the $s \times s$ matrices $A = (a_{ij})$, $\bar{A} = (a_{ij})$, and the s-dimensional vectors $b^T = (b_1, \ldots, b_s)$, $\bar{b}^T = (\bar{b}_1, \ldots, \bar{b}_s)$ and $c = (c_1, \ldots, c_s)^T$, the RKN scheme (10) can also be represented by the Butcher array

$$\begin{array}{c|c|c} c & \bar{A} & A \\ \hline & \bar{b}^T & b^T \end{array} \tag{11}$$

The method is said explicit if $a_{ij} = 0$ for $i \leq j$ and implicit otherwise.

If the coefficients A, b^T, c of the RKN method are equal to those of a Runge-Kutta method for first order ODE, then the Nyström method is said to be induced by this RK scheme and its coefficients satisfy

$$\bar{a}_{ij} = \sum_{k=1}^{s} a_{ik} a_{kj} \qquad \text{and} \qquad \bar{b}_i = \sum_{j=1}^{s} b_j a_{ji}. \tag{12}$$

Moreover, a RKN method is said to be a collocation scheme if it is obtained by applying collocation methods for first order differential equation (2).

We now investigate the properties a RKN scheme has to satisfy to be an orthogonal preserving scheme when applied to (1). We start with the following matrix differential system:

$$\ddot{Y}(t) = CY(t), \quad t > 0, \quad Y(0) = I_n, \quad \dot{Y}(0) = B, \tag{13}$$

where the matrix $C = B^2$ is symmetric seminegative definite and B is skew-symmetric.

Theorem 2. *The implicit RKN methods induced by the Gauss Legendre Runge Kutta schemes of order $2s$ are orthogonal integrators for differential systems (13).*

Proof. We give the proof for the Runge Kutta Nyström Gauss Legendre method with stage $s = 1$ and Butcher array

$$\begin{array}{c|c|c} 1/2 & 1/4 & 1/2 \\ \hline & 1/2 & 1 \end{array} \tag{14}$$

Applying the numerical scheme (14) to (13) we get:

$$\begin{aligned} K_1' &= C(Y_0 + \tfrac{h}{2}\dot{Y}_0 + \tfrac{h^2}{4}K_1'), \\ Y_1 &= Y_0 + h\dot{Y}_0 + \tfrac{1}{2}h^2 K_1', \\ \dot{Y}_1 &= \dot{Y}_0 + hK_1'. \end{aligned} \tag{15}$$

Hence

$$Y_1^T Y_1 = Y_0^T Y_0 + h(\dot{Y}^T Y_0 + Y_0^T \dot{Y}_0) + \tfrac{h^2}{2}(Y_0^T K_1' + {K_1'}^T Y_0 + 2\dot{Y}_0^T \dot{Y}_0) +$$

$$+ \tfrac{h^3}{2}(\dot{Y}_0^T K_1' + {K_1'}^T \dot{Y}_0) + \tfrac{h^4}{4}{K_1'}^T K_1'.$$

Substituting the expression of K_1' we get:

$$Y_1^T Y_1 = Y_0^T Y_0 + h(\dot{Y}^T Y_0 + Y_0^T \dot{Y}_0) + \tfrac{h^2}{2}(2Y_0^T C Y_0 + 2\dot{Y}_0^T \dot{Y}_0) +$$

$$+ \tfrac{h^3}{4}({Y_0}^T C\dot{Y}_0 + \dot{Y}_0^T C Y_0) + \tfrac{h^3}{2}(\dot{Y}_0^T K_1' + {K_1'}^T \dot{Y}_0) +$$

$$+ \tfrac{h^4}{4}(Y_0^T C K_1' + {K_1'}^T C Y_0 + {K_1'}^T K_1').$$

By the initial conditions and observing that $C = B^2$, the terms in h^2 and $\frac{h^3}{4}({Y_0}^T C\dot{Y}_0 + \dot{Y}_0^T C Y_0)$ vanish. Hence, substituting ricorsively the expression of K_1' all the powers of h vanish and so the result follows.

However, the positive result obtained by Theorem 2, are not still valid for more general second order nonlinear differential system.

Theorem 3. *The RKN Gauss Legendre schemes are not orthogonal integrators for the differential system (1).*

Proof. For the sake of simplicity we will give the proof in the linear nonautonomous case. Let us consider the differential system

$$\ddot{Y}(t) = C(t)Y(t), \quad t > 0, \quad Y(0) = I_n, \quad \dot{Y}(0) = B(0), \tag{16}$$

where $C(t) = \dot{B}(t) + B(t)B(t)$ and $B(t)$ is a skew-symmetric matrix. Applying the RKNGL scheme with $s = 1$ to (16), we get

$$\begin{aligned} K_1' &= C(\tfrac{h}{2})(Y_0 + \tfrac{1}{2}h\dot{Y}_0 + \tfrac{1}{4}h^2 K_1'), \\ Y_1 &= Y_0 + h\dot{Y}_0 + \tfrac{1}{2}h^2 K_1', \\ \dot{Y}_1 &= \dot{Y}_0 + hK_1'. \end{aligned} \tag{17}$$

Hence

$$Y_1^T Y_1 = Y_0^T Y_0 + h(\dot{Y}^T Y_0 + Y_0^T \dot{Y}_0) + \frac{h^2}{2}(Y_0^T C(\frac{h}{2})Y_0 + 2\dot{Y}_0^T \dot{Y}_0) + \dots.$$

Now, observe that the term in h^2, substituting the expression of $C(\frac{h}{2})$ and the values of Y_0 and of $\dot{Y}_0$, we get:

$$\frac{h^2}{2}\left(2\dot{B}(\frac{h}{2}) + 2\dot{B}(\frac{h}{2})\dot{B}(\frac{h}{2}) + 2\dot{B}(0)^T \dot{B}(0)\right), \tag{18}$$

which, in general, does not vanish and so $Y_1^T Y_1 = Y_0^T Y_0 + \mathcal{O}(h^2)$.

4 Non standard RKN schemes

In this section we generalize to RKN methods some approaches uses to solve orthogonal first order differential equations.

4.1 Projected RKN methods

Following the ideas proposed in [3] for first order ODEs on $\mathcal{O}_n$, we describe here a projected integrators for equation (1).

A projected methods consists of a two steps procedure:

- firstly, an approximation $\tilde{Y}_{k+1}$ of the solution of (1), provided by any explicit s-stage RKN method, is computed using (10);
- then, the QR factorization of $\tilde{Y}_{k+1}$ by the modified Gram-Schmidt process is performed, that is

$$\tilde{Y}_{k+1} = Q_{k+1}R_{k+1},$$

  and then the factor Q of the QR factorization is assumed as the approximation of the solution at t_{k+1}, i.e.,

$$Y_{k+1} = Q_{k+1}.$$

Proposition 1. *The projected RKN schemes preserve the order of accuracy of the RKN method they are based on.*

4.2 Cayley methods for second order ODEs

Another approach used for solving first order orthogonal systems is based on the transformation of the original system into a skew-symmetric one, obtained by continously applying the Cayley transform (see [4]).

Proposition 2. *If $Y(t)$ is an orthogonal matrix having for any t all eigenvalues different from -1, then there exists a unique smooth skew-symmetric matrix function $A(t)$ such that*

$$Y(t) = [I - A(t)]^{-1}[I + A(t)], \qquad t \geq 0. \tag{19}$$

The previous transformation is known as the Cayley transform of $Y(t)$.

An interesting remark is that for second order differential equations this approach does not lead to orthogonal schemes. In fact using (19), equation (1) can be transformed as in the following theorem.

Theorem 4. *Let $Y(t)$ be the solution of the differential systems (1), with all eigenvalue different from -1, for any t, then $A(t)$ given by (19) satisfies the second order differential system:*

$$\ddot{A} = H(A, \dot{A}), \qquad t > 0, \qquad A(0) = 0, \qquad \dot{A}(0) = \tfrac{1}{2}B, \tag{20}$$

where $H(A, \dot{A}) = \frac{1}{2}(I - A)C((I - A)^{-1}(I + A))(I - A) - 2\dot{A}(I - A)^{-1}\dot{A}$.

Proof. From (19) it follows that $(I - A)Y = (I + A)$; then differentiating twice

$$\dot{A} = \dot{Y} - \dot{A}Y - A\dot{Y}, \tag{21}$$

and

$$\ddot{A} = \ddot{Y} - \ddot{A}Y - 2\dot{A}\dot{Y} - A\ddot{Y}. \tag{22}$$

Hence $(I - A)\ddot{Y} - 2\dot{A}\dot{Y} = \ddot{A}(I + Y)$ and by (1)

$$\ddot{A} = (I - A)C(Y)(I + Y)^{-1} - 2\dot{A}\dot{Y}(I + Y)^{-1}.$$

Moreover, from (19) we also obtain

$$(I + Y)^{-1} = \frac{1}{2}(I - A), \tag{23}$$

and from (21)

$$\dot{Y} = 2(I - A)^{-1}\dot{A}(I - A)^{-1}. \tag{24}$$

Thus, substituting equality (23) and (24) into (22) and using the initial condition for equation (1) the statement follows.

Observe that a restarting procedure is required if there exists a τ such that $Y(\tau)$ has an eigenvalue equal to -1 (see [4]). (20).

Remark 2. We have to observe that the solution of (20) is not a skew-symmetric matrix function, because $H(\dot{A}, A)$ is not a skew-symmetric. Indeed, this result was expected, in fact from (21), $\dot{A}$ is not a curve on $o(n)$ and therefore, when we consider its derivative with respect to time, i.e. $\ddot{A}$, this does not belong to $o(n)$. Hence, the Cayley approach does not lead to orthogonal schemes.

5 Numerical Tests

In this section we present some numerical tests in order to illustrate the properties of the geodesics based methods. All the numerical results have been obtained by Matlab codes implemented on a scalar computer Alpha 200 5/433 with 512 Mb RAM. We compare the considered methods in terms of accuracy, deviation of the numerical solution from orthogonal structure and CPU time. The deviation from the orthogonal manifold is measured by $\|I_k - Y_k^T Y_k\|_F$, the accuracy by $\|Y(t_k) - Y_k\|_\infty$, where $\|\cdot\|_F$ and $\|\cdot\|_\infty$ denote respectively the Frobenius and the infinity norm on matrices and Y_k is the numerical approximation of the solution at the instant $t = t_k$.

Example 1. As first example we solve the constant linear second ordinary differential system

$$\ddot{Y} = CY, \quad Y(0) = I_2, \qquad \dot{Y}(0) = B,$$

Table 1. Example 1 performance at $T = 1$.

h	Method	Global error	Orthogonal error	CPU time
0.01	RKNGL1	0.0035	3.3813e-16	0.06
	PRKN2	0.0017	2.6037e-16	0.27
	RKN2	0.0018	3.1754e-4	0.21
	RKNGL2	2.4443e-7	3.7974e-16	0.11
	PRKN4	6.1915e-7	1.0245e-16	0.38
	RKN4	3.1272e-6	1.4040e-7	0.32
0.005	RKNGL1	8.6470e-4	3.6873e-16	0.11
	PRKN2	4.4099e-4	2.4999e-16	0.55
	RKN2	4.3600e-4	3.9733e-5	0.21
	RKNGL2	1.5280e-8	5.4897e-15	0.22
	PRKN4	3.8504e-8	5.0354e-16	0.72
	RKN4	1.9520e-7	4.3891e-9	0.75

where $B \begin{pmatrix} 0 & 1 & -3 & -4 \\ -1 & 0 & 2 & 2 \\ 3 & -2 & 0 & -3 \\ 4 & -2 & 3 & 0 \end{pmatrix}$ and $C = B^2$. Table 1 summerizes the results obtained solving the problem with constant step, on the interval $[0, 1]$. All the error are estimated at the final point of the integration interval.

As shown in Table (1) both projective and Gauss Legendre RKN schemes preserve the orthogonality with a machine accuracy. Furthermore, the direct application of an explicit RKN to the system provides an orthogonal error of the same order of the scheme.

Example 2. As second example we solve the nonautonomous second ordinary differential system

$$\ddot{Y} = \begin{pmatrix} -\sin^2(t) & \cos(t) \\ -\cos(t) & -\sin^2(t) \end{pmatrix} Y, \quad Y(0) = I_2, \qquad \dot{Y}(0) = 0,$$

whose solution is $Y(t) = \begin{pmatrix} \cos(1-\cos(t)) & \sin(1-\cos(t)) \\ -\sin(1-\cos(t)) & \cos(1-\cos(t)) \end{pmatrix}$, ([3]).

As proved in Lemma 3, for generally nonautonomous orthogonal second order systems, the collocation RKNGL schemes do not preserve the orthogonality of the solution.

Conclusion

With the aim of solving second ordinary differential systems preserving the orthogonal structure, we have investigated the properties of Runge Kutta Nyström Gauss Legendre methods. These schemes are orthogonal preserving only for linear constant second order ODEs. To tackle the problem, we have also proposed

Table 2. Example 2 performance at $T = 5$.

h	Method	Global error	Orthogonal error	CPU time
0.01	RKNGL1	1.4456e-5	4.5856e-5	0.27
	PRKN2	0.0018	0	0.50
	RKN2	0.0019	0.0032	0.47
	RKNGL2	1.8464e-10	3.4853e-10	0.50
	PRKN4	3.3782e-11	6.2803e-16	0.94
	RKN4	9.1295e-11	4.1622e-10	0.96
0.005	RKNGL1	3.6140e-6	1.0213e-5	0.43
	PRKN2	4.7764e-4	0	0.98
	RK2	3.9814e-4	0.0013	0.91
	RKNGL2	1.1528e-11	2.1871e-11	0.98
	PRKN4	2.1917e-12	0	1.86
	RK4	6.1917e-12	3.2724e-10	1.95

a semi-explicit projection procedure based on the Gram-Smith factorization and we have pointed out that the good performance of the Cayley approach for first order differential orthogonal systems are not showed for second order one.

For further research, we intend to extend the study of orthogonal behavior to other RKN schemes, as for instance symplectic RKN, and investigate the exponential map approach.

References

1. Calvo, M.P., Iserles, A., Zanna, A.: Runge-Kutta methods for orthogonal and isospectral flows, Appl. Numer. Math. 22 (1996), 153–164.
2. Del Buono, N., Lopez, L.: Runge Kutta type methods based on geodesics for systems of ODEs on the Stiefel manifold, BIT 41, 5, (2001) 912–923.
3. Dieci, L., Russell, D., Van Vleck, E. S.: Unitary integration and applications to continuous orthonormalization techniques, SIAM J. Num. Anal. 31 (1994), 261–281.
4. Diele, F., Lopez, L., Peluso, R.: The Cayley transform in the numerical solution of unitary differential systems , Ad. Comp. Math. 8 (1998), 317–334.
5. Edelman, A., Arias, T. A., Smith, S. T.: The geometry of algorithms with orthogonality constraints, SIAM J. Matrix Anal. Appl. 20 (1998), 303–353.
6. Hairer, E., Nørsett, S.P., Wanner, G.: Solving ordinary differential equations, Vol. I, Springer, Berlin, 1987.
7. Iserles, A., Nørsett, S.P.: On the solution of linear differential equation on Lie groups, Cambridge University Tech. Rep. DAMTP 1997/NA3.
8. Iserles, A., Munthe-Kaas, H. Z., Nørsett, S.P., Zanna, A.: Lie-group methods, Acta Numerica 9, CUP, (2000), pp. 215–365.
9. Petzold, L.R., Lay, L.O., Yen, J.: Numerical solution of highly oscillatory ordinary differential equations, Acta Numerica (1997), 437-483, Cambridge University Press.
10. Sanz-Serna, J.M., Calvo, M.P., Numerical Hamiltonian Problems, Chapman Hall, London, 1994.