BIT 34 (1994), 80–87.

THE NON-EXISTENCE OF SYMPLECTIC MULTI-DERIVATIVE RUNGE-KUTTA METHODS*

E. HAIRER[1], A. MURUA[2] and J. M. SANZ-SERNA[3]

[1] *Dept. de mathématiques, Université de Genève, Case postale 240 CH.1211 Genève 24, Switzerland, hairer@divsun.unige.ch*

[2] *Informatika Fakultatea, Euskal Herriko Unibertsitatea, 649 Postakutxa 20080 San Sebastian, Spain*

[3] *Departamento de Matemática Aplicada y Computación, Facultad de Ciencias Universidad de Valladolid, Valladolid, Spain, sanzserna@cpd.uva.es*

Abstract.

A sufficient condition for the symplecticness of q-derivative Runge-Kutta methods has been derived by F. M. Lasagni. In the present note we prove that this condition can only be satisfied for methods with $q \leq 1$, i.e., for standard Runge-Kutta methods. We further show that the conditions of Lasagni are also necessary for symplecticness so that no symplectic multi-derivative Runge-Kutta method can exist.

AMS subject classification: 65L06, 70H15.

Key words: Multi-derivative Runge-Kutta methods, symplectic methods, irreducible methods.

1. Introduction.

For the numerical solution of ordinary differential equations $y' = f(y)$ we consider *q-derivative Runge-Kutta methods* given by

$$y_1 = y_0 + \sum_{r=1}^{q} \sum_{i=1}^{s} b_i^{(r)} Y_i^{(r)}, \quad Y_i = y_0 + \sum_{r=1}^{q} \sum_{j=1}^{s} a_{ij}^{(r)} Y_j^{(r)} \tag{1.1}$$

where

$$Y_i^{(r)} = \frac{h^r}{r!}(D^r y)(Y_i) \quad \text{for } i = 1, \ldots, s \text{ and } r \geq 1. \tag{1.2}$$

Here, the differential operator D acts on functions Ψ: $\mathbb{R}^n \to \mathbb{R}^n$ and is defined by

---

* This research has been supported by project PB89-0351 (Dirección General de Investigación Científica y Técnica) and by project No. 20-32354.91 of Swiss National Science Foundation.

Received March 1993.

$$(D\Psi)(y) = \Psi'(y)f(y) \tag{1.3}$$

so that $D^1 y = f, D^2 y = f'f, D^3 y = f''(f,f) + f'f'f$, etc. For a detailed discussion of such methods we refer to Section II.13 of [4]. We closely follow the notation of this monograph and we shall use several results presented there.

In this paper we are interested in the numerical solution of *Hamiltonian systems*

$$p' = -\frac{\partial H}{\partial q}(p,q), \quad q' = \frac{\partial H}{\partial p}(p,q). \tag{1.4}$$

A characteristic property of the flow of (1.4) is that it is symplectic, i.e., it preserves the differential 2-form $dp \wedge dq$. Numerical methods, for which the numerical flow $(p_0, q_0) \mapsto (p_1, q_1)$ shares the same property are called *symplectic* (see [7] for an overview on symplectic integration.) Lasagni [6] was the first to study the symplecticness of multi-derivative Runge-Kutta methods. By a direct verification of the invariance of $dp \wedge dq$ he has shown that the conditions (we use the convention $b_i^{(m)} = a_{ij}^{(m)} = 0$ for $m > q$)

$$b_i^{(l)} b_j^{(r)} - b_i^{(l)} a_{ij}^{(r)} - b_j^{(r)} a_{ji}^{(l)} = \begin{cases} 0 & \text{for } i \neq j \\ b_i^{(l+r)} & \text{for } i = j \end{cases} \tag{1.5}$$

(for $i,j = 1, \ldots, s$; $l, r = 1, \ldots, q$) are sufficient for the method to be symplectic.

The main new results of this paper are the following: we first show (Section 2) that for irreducible methods the condition (1.5) can be satisfied only for the case $q = 1$. This is not very surprising because the number of conditions in (1.5), namely $qs(qs + 1)/2$, grows faster than the number of free parameters in the method which is $qs(s + 1)$.

In the second part of this paper (Section 3) we prove that for irreducible methods the condition (1.5) is not only sufficient but also necessary for symplecticness. Both results together imply that the method (1.1) cannot be symplectic unless for the case $q = 1$ which corresponds to standard Runge-Kutta methods.

2. Non-Existence of Methods Satisfying (1.5).

Methods, which have superfluous stages, have to be excluded in this section. Extending the notation of [2] to multi-derivative methods, we consider methods which are irreducible in the sense of the following definition.

DEFINITION. A multi-derivative method (1.1) is called *DJ-reducible* if for some non-empty index set $T \subset \{1, \ldots, s\}$

$$b_i^{(r)} = 0 \quad \text{for } i \in T \text{ and for } r \geq 1,$$

$$a_{ij}^{(r)} = 0 \quad \text{for } i \notin T, j \in T \text{ and for } r \geq 1.$$

Otherwise it is called *DJ-irreducible.*

The stage-values Y_i for $i \in T$ do not influence the numerical result y_1 and can be removed from the scheme.

THEOREM 1. *A DJ-irreducible multi-derivative Runge-Kutta method (1.1), which satisfies the condition (1.5) for all i, j and l, r, is equivalent to a standard Runge-Kutta method (i.e., $b_i^{(r)} = a_{ij}^{(r)} = 0$ for all i, j and $r \geq 2$).*

PROOF. Putting $T = \{i \mid b_i^{(1)} = 0\}$ we shall prove that
a) $b_i^{(r)} = 0$ for $i = 1, \ldots, s$ and $r \geq 2$;
b) $a_{ij}^{(r)} = 0$ for $i \notin T, j = 1, \ldots, s$ and $r \geq 2$;
c) $a_{ij}^{(1)} = 0$ for $i \notin T, j \in T$.
This implies $T = \emptyset$, otherwise the method would be DJ-reducible. Hence, by (a) and (b), all coefficients $b_i^{(r)}$, $a_{ij}^{(r)}$ vanish for $r \geq 2$ and the method is a standard Runge-Kutta method ($q = 1$).

For the proof of (a) we fix some $i \in \{1, \ldots, s\}$ and let m be the largest integer such that $b_i^{(m)} \neq 0$. We shall show that the assumption $m \geq 2$ leads to a contradiction. First, we put $j = i$ and $l = r = m$ in (1.5) to obtain $a_{ii}^{(m)} = b_i^{(m)}/2$. Next, we deduce from the same condition with $j = i$ and $r = m$ but arbitrary l that $a_{ii}^{(l)} = b_i^{(l)}/2$ for all $l \geq 1$. Finally, we put $l = 1$, $r = m - 1$ and obtain $b_i^{(m)} = b_i^{(1)} b_i^{(m-1)} - b_i^{(1)} a_{ii}^{(m-1)} - b_i^{(m-1)} a_{ii}^{(1)} = 0$, which contradicts the assumption $b_i^{(m)} \neq 0$.

Using (a) it follows from condition (1.5) with $l = 1$ and $r \geq 2$ (or $r = 1$ and $j \in T$) that $b_i^{(1)} a_{ij}^{(r)} = 0$. This implies $a_{ij}^{(r)} = 0$ for $i \notin T$ and $r \geq 2$ (or $i \notin T, j \in T$ and $r = 1$) and proves the statements (b) and (c). ■

3. Characterization of Symplectic Methods.

The aim of this section is to prove that for irreducible methods the condition (1.5) is also necessary for symplecticness. We introduce a kind of irreducibility and give an algebraic characterization for it. The equivalence of (1.5) with symplecticness is then obtained from a general result of [1] which is valid for methods whose solution can be represented as a B-series.

3.1. B-Series Representation of the Numerical Solution.

We consider the numerical solution y_1, given by (1.1), as a function of h and develop it into a Taylor series. This yields

$$y_1 = y_0 + \sum_{t \in T} \frac{h^{\rho(t)}}{\rho(t)!} \alpha(t) a(t) F(t)(y_0) \tag{3.1}$$

where T represents the set of rooted trees, $\rho(t)$ the number of vertices of the tree

t (order of t), $\alpha(t)$ is an integer coefficient, $F(t)(y_0)$ are the so-called elementary differentials and the coefficients $a: T \to \mathbb{R}$ depend on the parameters of the method. A series of the form (3.1) is called a *B-series*. Instead of discussing all the expressions appearing in (3.1) (see [4] for more details) we only collect those definitions which will be used in the sequel.

We denote by τ the only tree of order 1, and by

$$t = [t_1, \ldots, t_m]$$

the tree which consists of a root and of m leaving branches, to the end of which the trees $[t_1, \ldots, t_m]$ are attached. Similar to y_1 also the internal stages Y_i can be written in the form (3.1) with $a(t)$ replaced by $g_i(t)$. These coefficients are given by

$$a(t) = \sum_{r=1}^{q} \sum_{i=1}^{s} b_i^{(r)} g_i^{(r)}(t) \tag{3.2}$$

$$g_i(t) = \sum_{r=1}^{q} \sum_{j=1}^{s} a_{ij}^{(r)} g_j^{(r)}(t) \tag{3.3}$$

where $g_i^{(r)}(t)$ are the coefficients of the B-series for $Y_i^{(r)}$. For $t = \tau$ they are given by $g_i^{(1)}(\tau) = 1$, $g_i^{(r)}(\tau) = 0$ for $r \geq 2$, and for $t = [t_1, \ldots, t_m]$ we have the recursion (formula of Kastlunger; see exercise 2 of [4], Section II.13)

$$g_i^{(r)}(t) = \frac{\rho(t)}{r} \sum_{\substack{\lambda_1 + \ldots + \lambda_m = r-1 \\ \lambda_1, \ldots, \lambda_m \geq 0}} g_i^{(\lambda_1)}(t_1) \cdot \ldots \cdot g_i^{(\lambda_m)}(t_m) \tag{3.4}$$

where $g_i^{(0)}(t) = g_i(t)$ for all $t \in T$.

For later use we include the definition of the coefficients $\gamma(t)$: we put $\gamma(\tau) = 1$ and for $t = [t_1, \ldots, t_m]$ we define

$$\gamma(t) = \rho(t) \cdot \gamma(t_1) \cdot \ldots \cdot \gamma(t_m). \tag{3.5}$$

3.2. S-Irreducibility.

The following definition corresponds, in the case of standard Runge-Kutta methods, to the irreducibility introduced in [5].

DEFINITION. Two stages i and j of the method (1.1) are called *equivalent* if for every initial value problem $y' = f(y)$ and for every sufficiently small step size h it holds $Y_i = Y_j$. The method (1.1) is called *S-irreducible* if it possesses no equivalent stages.

For methods with equivalent stages the number of stages s can be reduced without changing the numerical solution y_1. The following lemma gives an algebraic characterization of S-irreducibility. It extends recent results of [1] and [3]. We shall assume that the set of trees T is ordered with τ being the first tree.

LEMMA 2. *For the q-derivative s-stage Runge-Kutta method (1.1) consider the* $q \cdot s \times \infty$ *matrix G whose columns are given by*

$$\frac{1}{\gamma(t)}(g_1^{(1)}(t), \ldots, g_s^{(1)}(t), \ldots, g_1^{(q)}(t), \ldots, g_s^{(q)}(t))^T \tag{3.6}$$

for $t \in T$. *Then, the method (1.1) is S-irreducible if and only if the matrix G has full rank* $q \cdot s$.

PROOF. If the stages i and j are equivalent it follows from the independency of the elementary differentials that the coefficients of the B-series for Y_i and Y_j are identical, i.e., $g_i(t) = g_j(t)$ for all $t \in T$. This implies that $g_i^{(1)}(t) = \rho(t) \cdot g_i(t_1) \cdot \ldots \cdot g_i(t_m)$ equals $g_j^{(1)}(t)$ for all t so that the rows i and j of G are the same.

For the proof of the "only if" part let us assume that the method is S-irreducible. The idea is to search for a matrix $C = (C_0, C_1, \ldots, C_{qs-1})$ with $C_l \in \mathbb{R}^\infty$ such that the product GC is a confluent Vandermonde type matrix. Putting $C_0 = (1, 0, 0, \ldots)^T$, the first column of GC is the vector (3.6) for $t = \tau$, i.e., $(1, 0, \ldots, 0)^T$ where the elements of $1 \in \mathbb{R}^s$ and $0 \in \mathbb{R}^s$ are all equal to 1 and 0, respectively. We will show in part (a) below that there exists a vector $C_1 \in \mathbb{R}^\infty$ (only finitely many components are non zero) such that the elements of

$$GC_1 = (\eta_1^{(1)}, \ldots, \eta_s^{(1)}, \eta_1^{(2)}, \ldots, \eta_s^{(2)}, \ldots, \eta_1^{(q)}, \ldots, \eta_s^{(q)})^T$$

satisfy: $\eta_1^{(1)}, \ldots, \eta_s^{(1)}$ are distinct and $\eta_1^{(2)}, \ldots, \eta_s^{(2)}$ are non zero. We then consider the polynomial $p(x)$ of degree $qs - 1$, defined by

$$p^{(j)}(i) = (j+1)! \cdot \eta_i^{(j+1)} \quad \text{for } i = 1, \ldots, s \text{ and } j = 0, \ldots, q-1. \tag{3.7}$$

In part (b) below we shall prove the existence of vectors $C_l (l \geq 2)$ such that

$$GC_l = \left(p^l(1), \ldots, p^l(s), \frac{1}{2!}\frac{d}{dx}p^l(1), \ldots, \frac{1}{2!}\frac{d}{dx}p^l(s), \ldots, \frac{1}{q!}\frac{d^{q-1}}{dx^{q-1}}p^l(1), \ldots, \frac{1}{q!}\frac{d^{q-1}}{dx^{q-1}}p^l(s)\right)^T. \tag{3.8}$$

In order to prove the linear independence of the vectors $GC_0, \ldots, GC_{qs-1}$ we suppose

$$\sum_{l=0}^{qs-1} d_l \cdot GC_l = 0 \tag{3.9}$$

and consider the polynomial $Q(y) = \sum_{l=0}^{qs-1} d_l y^l$. The condition (3.9) implies that the polynomial $Q(p(x))$ has zeros of multiplicity q at $1, 2, \ldots, s$. Since $p'(1), \ldots, p'(s)$ are non zero, $Q(y)$ must have zeros of multiplicity q at $p(1), \ldots, p(s)$. The fact that these zeros are distinct implies that $Q(y)$, a polynomial of degree $qs - 1$, vanishes identically so that $d_0 = \ldots = d_{qs-1} = 0$. Hence, the vectors $GC_0, \ldots, GC_{qs-1}$ are linearly independent and the "only if" statement of the lemma is proved.

a) Since the method is S-irreducible, there exists for every pair (i,j) with $i \neq j$ a tree $t \in T$ such that $g_i(t) \neq g_j(t)$. Due to the relation $g_i^{(1)}([t]) = \rho([t]) \cdot g_i(t)$, a suitable choice of trees t_j and coefficients c_j guarantees that the numbers

$$\eta_i = \sum_{j=1}^{k} c_j \frac{g_i^{(1)}([t_j])}{\gamma([t_j])}, \quad i = 1, \dots s \tag{3.10}$$

are distinct. We then put

$$\eta_i^{(l)} := \sum_{j=1}^{k} c_j \frac{g_i^{(l)}([t_j])}{\gamma([t_j])}, \quad i = 1, \dots, s;\ l = 1, \dots, q. \tag{3.11}$$

The values $\eta_i^{(1)} = \eta_i$ $(i = 1, \dots, s)$ are distinct. Furthermore, assuming that one tree among the t_j is τ, we can achieve $\eta_i^{(2)} \neq 0$ for all i by making the corresponding coefficient c_j large enough.

b) Using the definition (3.4) of $g_i^{(1)}(t)$ we see that the lth power of η_i is a linear combination of values $g_i^{(1)}(t)$, namely

$$\eta_i^l = \sum_{j_1, \dots, j_l = 1}^{k} c_{j_1} \cdot \ldots \cdot c_{j_l} \cdot \frac{g_i^{(1)}([t_{j_1}, \dots, t_{j_l}])}{\gamma([t_{j_1}, \dots, t_{j_l}])}.$$

Hence, there exists a vector C_l such that $\eta_1^l, \dots, \eta_s^l$ are the first s components of GC_l. We still have to prove that

$$\frac{d^r}{dx^r}(p^l(x))\big|_{x=i} = (r+1)! \sum_{j_1, \dots, j_l = 1}^{k} c_{j_1} \cdot \ldots \cdot c_{j_l} \cdot \frac{g_i^{(r+1)}([t_{j_1}, \dots, t_{j_l}])}{\gamma([t_{j_1}, \dots, t_{j_l}])}. \tag{3.12}$$

Using Kastlunger's formula (3.4) and the definitions (3.11) and (3.7), we see that the right-hand side of (3.12) can be written as

$$r! \sum_{j_1, \dots, j_l = 1}^{k} c_{j_1} \cdot \ldots \cdot c_{j_l} \sum_{\lambda_1 + \ldots + \lambda_l = r} \frac{g_i^{(\lambda_1+1)}([t_{j_1}])}{\gamma([t_{j_1}])} \cdot \ldots \cdot \frac{g_i^{(\lambda_l+1)}([t_{j_l}])}{\gamma([t_{j_l}])}$$

$$= \sum_{\lambda_1 + \ldots + \lambda_l = r} \binom{r}{\lambda_1, \dots, \lambda_l} p^{(\lambda_1)}(i) \cdot \ldots \cdot p^{(\lambda_l)}(i).$$

However, Leibniz' rule implies that this expression is equal to the left-hand side of (3.12). ■

3.3. *Necessary and Sufficient Condition for Symplecticness.*

In the article [1] it is shown that the B-series (3.1) defines a symplectic transformation for each problem of the form (1.4) if and only if the coefficients $a(t)$ satisfy the relation

$$\frac{a(u \circ v)}{\gamma(u \circ v)} + \frac{a(v \circ u)}{\gamma(v \circ u)} = \frac{a(u)}{\gamma(u)} \cdot \frac{a(v)}{\gamma(v)} \quad \text{for all } u, v \in T. \tag{3.13}$$

Here we have used the notation

$$u \circ v = [u_1, \ldots, u_m, v], \quad v \circ u = [v_1, \ldots, v_l, u]$$

for $u = [u_1, \ldots, u_m]$ and $v = [v_1, \ldots, v_l]$. The coefficients $\gamma(t)$ are those of (3.5). We shall apply this criterion to the coefficients (3.2) of our mult-derivative method (1.1).

LEMMA 3. *Let be*

$$m_{ij}^{(l,r)} := b_i^{(l)} b_j^{(r)} - b_i^{(l)} a_{ij}^{(r)} - b_j^{(r)} a_{ji}^{(l)} - \delta_{ij} b_i^{(l+r)} \tag{3.14}$$

where δ_{ij} is the Kronecker delta. The coefficients $a(t)$ of (3.2) then satisfy

$$\frac{a(u \circ v)}{\gamma(u \circ v)} + \frac{a(v \circ u)}{\gamma(v \circ u)} - \frac{a(u)}{\gamma(u)} \cdot \frac{a(v)}{\gamma(v)} = - \sum_{l,r=1}^{q} \sum_{i,j}^{s} m_{ij}^{(l,r)} \cdot \frac{g_i^{(l)}(u)}{\gamma(u)} \frac{g_j^{(r)}(v)}{\gamma(v)}$$

for all trees $u, v \in T$.

PROOF. We multiply (3.14) by $g_i^{(l)}(u)/\gamma(u)$ and $g_j^{(r)}(v)/\gamma(v)$, sum over all appearing indices and insert formula (3.3). In this way we obtain

$$\begin{aligned} \sum_{l,r=1}^{q} \sum_{i,j}^{s} m_{ij}^{(l,r)} \cdot \frac{g_i^{(l)}(u)}{\gamma(u)} \frac{g_j^{(r)}(v)}{\gamma(v)} &= \left(\sum_{l=1}^{q} \sum_{i=1}^{s} b_i^{(l)} \frac{g_i^{(l)}(u)}{\gamma(u)} \right) \left(\sum_{r=1}^{q} \sum_{j=1}^{s} b_j^{(r)} \frac{g_j^{(r)}(v)}{\gamma(v)} \right) \\ - \sum_{l=1}^{q} \sum_{i=1}^{s} b_i^{(l)} \frac{g_i^{(l)}(u)}{\gamma(u)} \frac{g_i(v)}{\gamma(v)} &- \sum_{r=1}^{q} \sum_{j=1}^{s} b_j^{(r)} \frac{g_j^{(r)}(v)}{\gamma(v)} \frac{g_j(u)}{\gamma(u)} - \sum_{l,r=1}^{q} \sum_{i=1}^{s} b_i^{(l+r)} \frac{g_i^{(l)}(u)}{\gamma(u)} \frac{g_i^{(r)}(v)}{\gamma(v)}. \end{aligned} \tag{3.15}$$

On the other hand, it follows from (3.4) and (3.5) that

$$\frac{g_i^{(r)}(u \circ v)}{\gamma(u \circ v)} = \sum_{k=0}^{r-1} \frac{r-k}{r} \cdot \frac{g_i^{(r-k)}(u)}{\gamma(u)} \frac{g_i^{(k)}(v)}{\gamma(v)}$$

so that

$$\frac{g_i^{(r)}(u \circ v)}{\gamma(u \circ v)} + \frac{g_i^{(r)}(v \circ u)}{\gamma(v \circ u)} = \frac{g_i^{(r)}(u)}{\gamma(u)} \frac{g_i(v)}{\gamma(v)} + \frac{g_i^{(r)}(v)}{\gamma(v)} \frac{g_i(u)}{\gamma(u)} + \sum_{k=1}^{r-1} \frac{g_i^{(r-k)}(u)}{\gamma(u)} \frac{g_i^{(k)}(v)}{\gamma(v)}. \tag{3.16}$$

If we multiply (3.16) by $b_i^{(r)}$ and sum over all i and r we obtain exactly the last three expressions of (3.15). The definition (3.2) of the coefficients $a(t)$ thus yields the statement of the lemma. ■

It follows from Lemma 3 that the condition (3.13) holds if and only if $G^T M G = 0$, where G is the matrix of Lemma 2 and

$$M = \begin{pmatrix} (m_{ij}^{(1,1)})_{i,j} & \cdots & (m_{ij}^{(1,q)})_{i,j} \\ \vdots & & \vdots \\ (m_{ij}^{(q,1)})_{i,j} & \cdots & (m_{ij}^{(q,q)})_{i,j} \end{pmatrix}. \tag{3.17}$$

Since G has full rank for S-irreducible methods, this is equivalent to $M = 0$ which is identical to condition (1.5). We thus have proved the following result.

THEOREM 4. *An S-irreducible multi-derivative Runge-Kutta method is symplectic if and only if condition (1.5) is satisfied.*

Combining the statements of Theorems 1 and 4 we obtain the main result of this article.

THEOREM 5. *A multi-derivative Runge-Kutta method, which is DJ- and S-irreducible, cannot be symplectic unless* $b_i^{(r)} = a_{ij}^{(r)} = 0$ *for all* i, j *and* $r \geq 2$.

REFERENCES

1. M. P. Calvo and J. M. Sanz-Serna, *Canonical B-series.* To appear in Numer. Math.
2. G. Dahlquist and R. Jeltsch, *Generalized disks of contractivity for explicit and implicit Runge-Kutta methods.* TRITA-NA Report 7906 (1979).
3. E. Hairer, *Backward analysis of numerical and symplectic methods.* To appear in Advances in Comput. Math.
4. E. Hairer, S. P. Nørsett and G. Wanner, *Solving Ordinary Differential Equations I. Nonstiff Problems.* Second Edition. Springer-Verlag, Berlin, 1993.
5. W. H. Hundsdorfer and M. N. Spijker, *A note on B-stability of Runge-Kutta methods.* Numer. Math. 36 (1981), p. 319–331.
6. F. M. Lasagni, *Integration methods for Hamiltonian differential equations.* Unpublished Manuscript.
7. J. M. Sanz-Serna, *Symplectic integrators for Hamiltonian problems: an overview.* Acta Numerica 1 (1992), p. 243–286.