Numer. Math. 67: 161–175 (1994)

Numerische Mathematik

Canonical B-series

M.P. Calvo, J.M. Sanz-Serna

Departamento de Matemática Aplicada y Computación, Facultad de Ciencias, Universidad de Valladolid, Valladolid, Spain

Received March 15, 1993

Summary. B-series provide a powerful general tool to express numerical methods for differential equations. Many differential equations are of Hamiltonian form and there has been much recent interest in constructing so-called canonical or symplectic integrators for the Hamiltonian case. In this paper we provide a necessary and sufficient condition for a B-series to correspond to a canonical method.

Mathematics Subject Classification (1991): 65L20

1. Introduction

Given a system of ordinary differential equations in $\mathbb{R}^D$

$$\frac{d\mathbf{y}}{dt} = \mathbf{f}(\mathbf{y}), \tag{1.1}$$

a B-series [4], [3] is a formal expression

$$\mathbf{y} + \sum_{n=1}^{\infty} \frac{h^n}{n!} \sum_{\rho\tau \in RT_n} c(\rho\tau)\mathbf{F}(\rho\tau)(\mathbf{y}), \tag{1.2}$$

where h is a real parameter, RT_n is the set of all rooted trees $\rho\tau$ with n vertices, $c(\rho\tau)$ is a real coefficient associated with $\rho\tau$ and $\mathbf{F}(\rho\tau)(\mathbf{y})$ is the elementary differential corresponding to $\rho\tau$ evaluated at $\mathbf{y} \in \mathbb{R}^D$. (The notions of rooted tree and elementary differential are revised in Sects. 2 and 3 below.) B-series more general than (1.2) are possible: the series may begin with a term $a\mathbf{y}$, a a constant, rather than with $\mathbf{y}$. However in this paper we are only concerned with the format (1.2).

B-series are a powerful tool for studying numerical methods for the integration of (1.1), see e.g. [3]. Assume that (1.1) is integrated by a Runge-Kutta (RK) or by a q-derivative ($q \geq 2$) Runge-Kutta (qRK) method. Denote by $\mathbf{y}^* = \psi_{h,\mathbf{f}}(\mathbf{y})$ the result of a step of length h starting from $\mathbf{y}$. Then the formal expansion of $\mathbf{y}^*$ in powers of h is of the form (1.2); the coefficients $c(\rho\tau)$ depend on the specific method, but not on the problem (1.1) being solved. Furthermore, let us now denote by $\mathbf{y}^* = \varphi_{h,\mathbf{f}}(\mathbf{y})$

the true value at time $t = h$ of the solution of (1.1) with initial value $\mathbf{y}$ at $t = 0$ (i.e., the mapping $\varphi_{h,\mathbf{f}}$ is the h-flow of (1.1)). Then the formal expansion of $\mathbf{y}^* = \varphi_{h,\mathbf{f}}(\mathbf{y})$ in powers of h is also a B-series: the B-series such that, for each $\rho\tau$,

$$c(\rho\tau) = \alpha(\rho\tau), \tag{1.3}$$

where $\alpha(\rho\tau)$ denotes the number of monotonic labellings of $\rho\tau$.

There has been much interest (see [7] for a survey) in developing numerical methods adapted to the special case where (1.1) is a Hamiltonian system, i.e., $D = 2d$ (d is called the number of degrees of freedom) and

$$\mathbf{f}(\mathbf{y}) = \Xi \nabla H(\mathbf{y}), \tag{1.4}$$

where $\Xi = -\Xi^{\mathrm{T}} = -\Xi^{-1}$ is the matrix

$$\Xi = \begin{bmatrix} 0_d & I_d \\ -I_d & 0_d \end{bmatrix}, \tag{1.5}$$

∇ denotes the gradient operator $\nabla = (\partial/\partial y^1, \ldots, \partial/\partial y^{2d})$, and H is a real valued function (the Hamiltonian). Of particular interest are so-called symplectic (or canonical) methods. A mapping $\psi : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ is called symplectic if

$$\psi'^{\mathrm{T}} \Xi \psi' \equiv \Xi, \tag{1.6}$$

where ψ' is the Jacobian matrix of ψ. For each H and each h, the true flow $\varphi_{h,\Xi\nabla H}$ of a Hamiltonian system is a symplectic transformation; this property is the most important feature of Hamiltonian flows. A numerical method ψ is called symplectic if $\psi_{h,\Xi\nabla H}$ is a symplectic transformation for each step-size h and each Hamiltonian H.

It is then of clear interest to ascertain under which conditions on the coefficients $c(\rho\tau)$, the B-series (1.2) defines a symplectic transformation for each h and each $\mathbf{f}$ of the form (1.4). The main result of this paper (Sect. 2) answers this question. It turns out that if a B-series satisfies the necessary and sufficient condition for symplecticness, then the corresponding order conditions are greatly simplified (a fact that was known [8] in the particular case where the B-series arises from a symplectic RK method). The main result is proved in Sects. 3–5. The final Sect. 6 illustrates the application of the main result to the particular instance of RK methods.

2. The main result

It is well known [2], [3] that the conditions for an RK or qRK method to have order $\geq r$ are written invoking rooted trees $\rho\tau$ of order $n(\rho\tau) \leq r$ (i.e., having r or fewer vertices). The left section of Fig. 1 depicts the rooted trees of order ≤ 4. The root of each rooted tree $\rho\tau$ has been highlighted by appending a cross.

If in $\rho\tau_{3,1}$ and $\rho\tau_{3,2}$ in the figure we disregard the location of the roots, then both graphs are identical; they consist of the same vertices joined by the same set of edges. The graph obtained by disregarding the location of the root in a rooted tree is called a free tree, or simply, a tree. Thus a tree τ can be seen as an equivalence class of rooted trees. On the right of Fig. 1 we have displayed the trees τ of order ≤ 4. A tree and the rooted trees belonging to it appear in the same row.

Now choose a tree τ of order ≥ 2 and a pair of adjacent vertices i and j in τ. By choosing i (resp. j) to play the role of root, we obtain a rooted tree $\rho\tau_i$ (resp. $\rho\tau_j$).

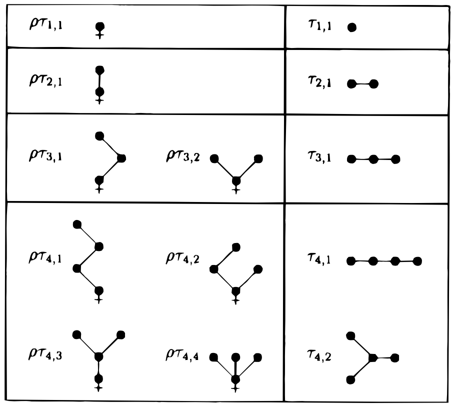

Fig. 1. Rooted n-trees and n-trees, $n = 1, 2, 3, 4$

We say that the rooted trees $\rho\tau_i$ and $\rho\tau_j$ are *neighbours*. Thus, in $\tau_{3,1}$ if we choose the leftmost vertex to be the root we obtain $\rho\tau_{3,1}$, if we choose the central vertex to be the root we obtain $\rho\tau_{3,2}$; therefore $\rho\tau_{3,1}$ and $\rho\tau_{3,2}$ are neighbours. In a similar manner, $\rho\tau_{4,1}$ and $\rho\tau_{4,2}$ are neighbours; $\rho\tau_{4,3}$ and $\rho\tau_{4,4}$ are neighbours. The 'straight' tree with 5 vertices $a - b - c - d - e$ comprises three rooted trees. The rooted tree obtained by setting the root at a is not a neighbour of the rooted tree obtained by setting the root at c.

Figure 2 contains a tree τ, two neighbours $\rho\tau_i$ and $\rho\tau_j$ and also the rooted trees $\rho\tau_I$, $\rho\tau_J$, with roots at i and j respectively, that arise when the edge joining i and j is removed from τ.

We may now give the main result in this paper. (For each $\rho\tau$, $\gamma(\rho\tau)$ represents, as in [2] or [3], the corresponding density.)

Theorem 2.1. *The B-series (1.2) is canonical, i.e., defines a symplectic transformation for each h and each Hamiltonian problem (1.1), (1.4), if, and only if, for each pair of neighbours $\rho\tau_i$ and $\rho\tau_j$*

$$\text{(2.1)} \qquad \frac{c(\rho\tau_i)}{\alpha(\rho\tau_i)\gamma(\rho\tau_i)} + \frac{c(\rho\tau_j)}{\alpha(\rho\tau_j)\gamma(\rho\tau_j)} = \frac{c(\rho\tau_I)}{\alpha(\rho\tau_I)\gamma(\rho\tau_I)}\,\frac{c(\rho\tau_J)}{\alpha(\rho\tau_J)\gamma(\rho\tau_J)}.$$

The condition in (2.1) can be rewritten in a slightly different form. Denote by $[\rho\tau]$ the root tree obtained by grafting a rooted tree $\rho\tau$ into a new root and denote by $\rho\tau_1 \cdot \rho\tau_2$ the rooted tree obtained from $\rho\tau_1$, $\rho\tau_2$ by identifying their roots. Then in (2.1), $\rho\tau_i = \rho\tau_I \cdot [\rho\tau_J]$ and $\rho\tau_j = [\rho\tau_I] \cdot \rho\tau_J$ and the necessary and sufficient condition for canonicity is that for any pair $\rho\tau_I$, $\rho\tau_J$ of rooted trees

$$\frac{c(\rho\tau_I \cdot [\rho\tau_J])}{\alpha(\rho\tau_I \cdot [\rho\tau_J])\gamma(\rho\tau_I \cdot [\rho\tau_J])} + \frac{c([\rho\tau_I] \cdot \rho\tau_J)}{\alpha([\rho\tau_I] \cdot \rho\tau_J)\gamma([\rho\tau_I] \cdot \rho\tau_J)}$$
$$= \frac{c(\rho\tau_I)}{\alpha(\rho\tau_I)\gamma(\rho\tau_I)}\,\frac{c(\rho\tau_J)}{\alpha(\rho\tau_J)\gamma(\rho\tau_J)}.$$

Before we prove the main theorem in Sects. 3–5, let us point out an important implication. Assume that a canonical B-series with coefficients $c(\rho\tau)$ has order of consistency $\geq r - 1$, $r > 2$, i.e. that it differs from the B-series of the true flow

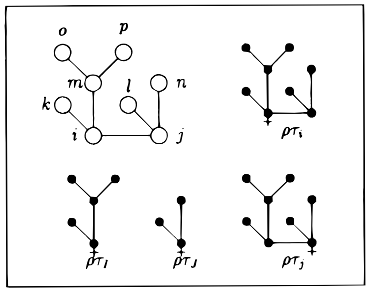

Fig. 2. The construction for the main result

in terms $O(h^r)$. Then, according to (1.3), $c(\rho\tau)/\alpha(\rho\tau) = 1$ whenever $\rho\tau$ has order $\leq r-1$. Consider then a pair of neighbours $\rho\tau_i$, $\rho\tau_j$ of order r and apply the condition (2.1) to obtain

$$\frac{c(\rho\tau_i)}{\alpha(\rho\tau_i)\gamma(\rho\tau_i)} + \frac{c(\rho\tau_j)}{\alpha(\rho\tau_j)\gamma(\rho\tau_j)} = \frac{1}{\gamma(\rho\tau_I)}\ \frac{1}{\gamma(\rho\tau_J)}.$$

The right hand side of this equality is easily seen [8] to equal $\gamma(\rho\tau_i)^{-1} + \gamma(\rho\tau_j)^{-1}$, and therefore

$$\frac{c(\rho\tau_i)}{\alpha(\rho\tau_i)\gamma(\rho\tau_i)} + \frac{c(\rho\tau_j)}{\alpha(\rho\tau_j)\gamma(\rho\tau_j)} = \frac{1}{\gamma(\rho\tau_i)} + \frac{1}{\gamma(\rho\tau_j)}. \tag{2.2}$$

Now, if the order condition $c(\rho\tau_i)/\alpha(\rho\tau_i) = 1$ associated with $\rho\tau_i$ holds, then (2.2) shows that $c(\rho\tau_j)/\alpha(\rho\tau_j) = 1$, i.e., the order condition associated with $\rho\tau_j$ holds. We have proved the following lemma:

Lemma 2.1. *For a canonical B-series of order $\geq r-1$, $r \geq 2$, the order conditions corresponding to two neighbours $\rho\tau_i$, $\rho\tau_j$ of order r are equivalent.*

There are two ways in which this result can be strengthened. The first is to note that a rooted tree may be its own neighbour. This happens, for instance in the tree $\tau_{2,1}$: setting the root at the left vertex leads to $\rho\tau_{2,1}$, and setting the root at the adjacent, right vertex also leads to $\rho\tau_{2,1}$, so that $\rho\tau_{2,1}$ is its own neighbour.

If $\rho\tau_i = \rho\tau_j$ and the conditions of the lemma hold, then (2.2) yields $c(\rho\tau_i)/\alpha(\rho\tau_i) = 1$, so that the order condition for $\rho\tau_i$ is automatically satisfied. Trees, such as $\tau_{2,1}$ or $\tau_{4,1}$, that contain a rooted tree $\rho\tau_i$ that is its own neighbour were called in [8] superfluous.

The second way in which the lemma may be strengthened is to note that, given two rooted trees $\rho\tau_i$ and $\rho\tau_k$ in the same tree, there is a chain $\rho\tau_i = \rho\tau_{i_1}, \rho\tau_{i_2}, \rho\tau_{i_3}, \ldots, \rho\tau_{i_l} = \rho\tau_k$ where two consecutive links are neighbours. Hence, under the conditions of the lemma, the order conditions corresponding to any two rooted trees $\rho\tau_i$, $\rho\tau_k$ in the same tree are equivalent. We summarize our discussion as follows:

Theorem 2.2. *For a canonical B-series of order $\geq r-1$, $r \geq 2$ to have order of consistency $\geq r$, it is (necessary and) sufficient that for each nonsuperfluous tree τ with r vertices there exists a rooted tree $\rho\tau$ in τ for which the order condition $c(\rho\tau) = \alpha(\rho\tau)$ holds.*

3. Preliminary results

In this section we present some results on graph theory that are needed to prove the main Theorem 2.1.

For each rooted tree $\rho\tau$, we denote by $\Sigma(\rho\tau)$ the group of its symmetries ([2], Definition 140F). The number of elements in $\Sigma(\rho\tau)$ is denoted, as usual, by $\sigma(\rho\tau)$. Between $\sigma(\rho\tau)$ and the quantities $n(\rho\tau)$, $\alpha(\rho\tau)$ and $\gamma(\rho\tau)$ introduced before there is a relation ([2], Theorem 145E)

$$n(\rho\tau)! = \alpha(\rho\tau)\gamma(\rho\tau)\sigma(\rho\tau). \tag{3.1}$$

Given a rooted tree $\rho\tau$, we consider the following equivalence relation. We say that two vertices v and w in $\rho\tau$ are equivalent if there exists a symmetry $S \in \Sigma(\rho\tau)$ such that $S(v) = w$. We denote by $S(v, \rho\tau)$ the number of elements of the equivalence class of the vertex v.

Following a standard notation, the symbol $[\rho\tau_1^{m_1}, \rho\tau_2^{m_2}, \dots, \rho\tau_k^{m_k}]$ refers to the rooted tree of order $1 + \sum_{i=1}^k m_i\, n(\rho\tau_i)$ obtained by grafting to a common root m_1 copies of the rooted tree $\rho\tau_1$, m_2 copies of the rooted tree $\rho\tau_2$, etc... The following result is (almost) evident.

Lemma 3.1. *Let $\rho\tau = [\rho\tau_1^{m_1}, \rho\tau_2^{m_2}, \dots, \rho\tau_k^{m_k}]$ with $\rho\tau_i$ pairwise distinct and $m_i \geq 1$. Then for each vertex v of $\rho\tau_1$*

$$S(v, \rho\tau) = m_1 S(v, \rho\tau_1)$$

Let us now consider four rooted trees $\rho\tau_i$, $\rho\tau_j$, $\rho\tau_I$, $\rho\tau_J$ as those described in connection with Theorem 2.1 (see Fig. 2).

Lemma 3.2. *With $\rho\tau_i$, $\rho\tau_j$, $\rho\tau_I$, $\rho\tau_J$ as above, let w be a vertex in $\rho\tau_J$. Then*

$$\frac{S(w, \rho\tau_i)}{\sigma(\rho\tau_i)} = \frac{S(i, \rho\tau_I)}{\sigma(\rho\tau_I)} \frac{S(w, \rho\tau_J)}{\sigma(\rho\tau_J)}$$

(i is the root of $\rho\tau_I$).

Proof. We begin by noticing that clearly $S(i, \rho\tau_I) = 1$; this factor is only included in the formula for analogy with expressions to be found later. Assume that

$$\rho\tau_i = [\rho\tau_J^{m_0}, \rho\tau_1^{m_1}, \dots, \rho\tau_k^{m_k}],$$

with distinct $\rho\tau_J$, $\rho\tau_1$, ..., $\rho\tau_k$ and positive $m_0, \dots, m_k$. Then

$$\rho\tau_I = [\rho\tau_J^{m_0-1}, \rho\tau_1^{m_1}, \dots, \rho\tau_k^{m_k}].$$

By standard results ([2], Theorem 144A)

$$\begin{aligned} \sigma(\rho\tau_i) &= m_0!\sigma(\rho\tau_J)^{m_0} m_1!\sigma(\rho\tau_1)^{m_1} \cdots m_k!\sigma(\rho\tau_k)^{m_k}, \\ \sigma(\rho\tau_I) &= (m_0 - 1)!\sigma(\rho\tau_J)^{m_0-1} m_1!\sigma(\rho\tau_1)^{m_1} \cdots m_k!\sigma(\rho\tau_k)^{m_k}; \end{aligned}$$

therefore $\sigma(\rho\tau_i) = m_0\sigma(\rho\tau_I)\sigma(\rho\tau_J)$. On the other hand by Lemma 3.1

$$S(w, \rho\tau_i) = m_0 S(w, \rho\tau_J).$$

This completes the proof. □

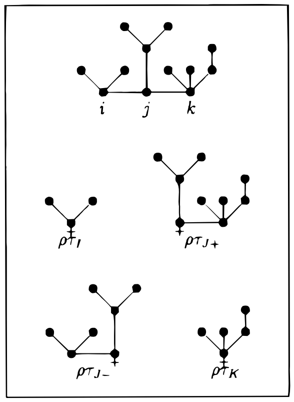

Fig. 3. The construction for Lemma 3.3

Next we consider a more complicated situation. Let τ be a tree of order ≥ 3 and consider distinct vertices i, j, k in τ with i and k adjacent to j (Fig. 3). Denote by $\rho\tau_I$, $\rho\tau_{J+}$ the rooted trees obtained by removing the edge i–j from τ. Denote by $\rho\tau_{J-}$, $\rho\tau_K$ the rooted trees obtained by removing the edge j–k from τ.

Lemma 3.3. *With the notations just introduced, let v (resp. w) be a vertex in $\rho\tau_I$ (resp. $\rho\tau_K$). Then*

$$\frac{S(v,\rho\tau_I)S(w,\rho\tau_{J+})}{\sigma(\rho\tau_I)\sigma(\rho\tau_{J+})} = \frac{S(v,\rho\tau_{J-})S(w,\rho\tau_K)}{\sigma(\rho\tau_{J-})\sigma(\rho\tau_K)}.$$

Proof. The proof is very similar to that of Lemma 3.2 and will not be given. □

Now (Fig. 4) consider a tree τ and a chain of adjacent vertices i_1, i_2, ..., i_k, $k \geq 1$. Denote by $\rho\tau_{i_1}$ (resp. $\rho\tau_{i_k}$) the rooted tree obtained by choosing i_1 (resp. i_k) to be the root. If $k \geq 2$ denote by $\rho\tau_{I_j-}$ and $\rho\tau_{I_{j+1}+}$, $j = 1, \ldots, k-1$ the rooted trees obtained by removing the edge i_j–i_{j+1} from τ.

Lemma 3.4. *With the preceding notation*

$$\begin{aligned}
\frac{S(i_k,\rho\tau_{i_1})}{\sigma(\rho\tau_{i_1})} &= \frac{S(i_1,\rho\tau_{I_1-})}{\sigma(\rho\tau_{I_1-})}\frac{S(i_k,\rho\tau_{I_2+})}{\sigma(\rho\tau_{I_2+})} \\
&= \frac{S(i_1,\rho\tau_{I_2-})}{\sigma(\rho\tau_{I_2-})}\frac{S(i_k,\rho\tau_{I_3+})}{\sigma(\rho\tau_{I_3+})} \\
&= \ldots\ldots\ldots\ldots\ldots \\
&= \frac{S(i_1,\rho\tau_{i_k})}{\sigma(\rho\tau_{i_k})}.
\end{aligned}$$

Proof. If $k = 1$ the expression above contains only one equality, which is obviously satisfied. If $k = 2$, there are two equalities to prove; both are implied by Lemma 3.2. For $k \geq 3$, there are k equalities; the first and last have been proved in Lemma 3.2 above. The second is a consequence of Lemma 3.3 with i_1, i_2, i_3 playing the role of i, j, k respectively, etc...

The final result in this section will be the key point in the proof of the main theorem.

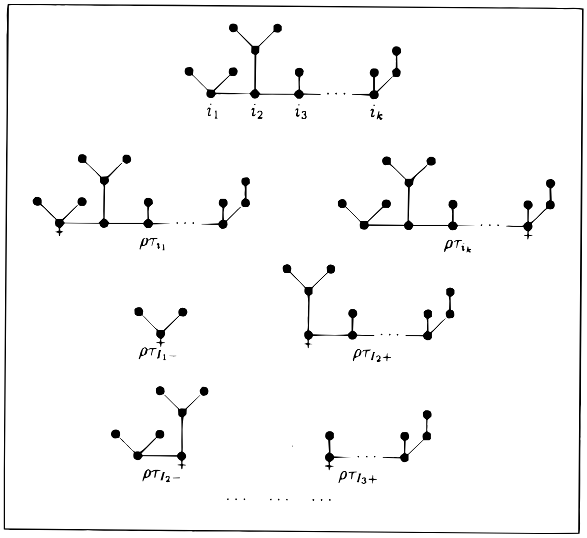

Fig. 4. The construction for Lemma 3.4

Lemma 3.5. *In the situation of Lemma 3.4, suppose that the B-series (1.2) satisfies the canonicity condition in Theorem 2.1. Then*

$$\begin{aligned} &- \frac{c(\rho\tau_{i_1})S(i_k,\rho\tau_{i_1})}{n(\rho\tau_{i_1})!} + \frac{c(\rho\tau_{I_1-})S(i_1,\rho\tau_{I_1-})}{n(\rho\tau_{I_1-})!}\frac{c(\rho\tau_{I_2+})S(i_k,\rho\tau_{I_2+})}{n(\rho\tau_{I_2+})!} \\ &- \frac{c(\rho\tau_{I_2-})S(i_1,\rho\tau_{I_2-})}{n(\rho\tau_{I_2-})!}\frac{c(\rho\tau_{I_3+})S(i_k,\rho\tau_{I_3+})}{n(\rho\tau_{I_3+})!} \\ &+ \cdots + (-1)^{k+1}\frac{c(\rho\tau_{i_k})S(i_1,\rho\tau_{i_k})}{n(\rho\tau_{i_k})!} = 0. \end{aligned}$$

Proof. By using (3.1) and Lemma 3.4, we conclude that we have to prove that

$$\begin{aligned} &- \frac{c(\rho\tau_{i_1})}{\alpha(\rho\tau_{i_1})\gamma(\rho\tau_{i_1})} + \frac{c(\rho\tau_{I_1-})}{\alpha(\rho\tau_{I_1-})\gamma(\rho\tau_{I_1-})}\frac{c(\rho\tau_{I_2+})}{\alpha(\rho\tau_{I_2+})\gamma(\rho\tau_{I_2+})} \\ &- \frac{c(\rho\tau_{I_2-})}{\alpha(\rho\tau_{I_2-})\gamma(\rho\tau_{I_2-})}\frac{c(\rho\tau_{I_3+})}{\alpha(\rho\tau_{I_3+})\gamma(\rho\tau_{I_3+})} \\ &+ \cdots + (-1)^{k+1}\frac{c(\rho\tau_{i_k})}{\alpha(\rho\tau_{i_k})\gamma(\rho\tau_{i_k})} = 0. \end{aligned} \tag{3.2}$$

Now we use (2.1) $k-1$ times, successively taking the vertices i_α, $i_{\alpha+1}$, $\alpha = 1,\ldots,k-1$ to play the roles of i and j in (2.1). This easily leads to (3.2). □

4. Proof of the main result: sufficiency

4.1. Elementary differentials for Hamiltonian problems

Let us begin by considering the elementary differentials $\mathbf{F}(\rho\tau)(\mathbf{y})$ in (1.2). For the rooted tree $\rho\tau_{3,2}$ on the left of Fig. 5 this is the vector whose i-th component is given

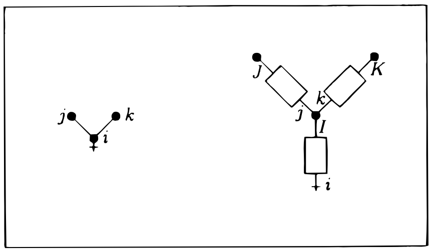

Fig. 5. Standard representation of a rooted tree and the representation of the same rooted tree useful in the Hamiltonian case

by

$$f^i_{jk} f^j f^k; \tag{4.1}$$

here and elsewhere we use the convention of summation on repeated indices, superscripts denote components and subscripts derivatives (e.g., f^j_k is the partial derivative of the j-th component of **f** with respect to the k-th component y^k of **y**). In (4.1) the functions are evaluated at **y**. The general rule is that there is an index i, j, k, $\dots$ per vertex and that a vertex k with sons $l_1, \dots, l_m$ introduces a factor $f^k_{l_1,\dots,l_m}$.

In the particular case where (1.4) holds and the system is Hamiltonian, it is useful to write the elementary differentials in terms of H rather than in terms of **f**. If ξ_{ij} are the elements of the matrix Ξ in (1.5), then (4.1) becomes in the Hamiltonian case

$$\xi_{iI} H_{Ijk} \xi_{jJ} H_J \xi_{kK} H_K . \tag{4.2}$$

This expression may be easily remembered after drawing the corresponding rooted tree $\rho\tau_{3,2}$ in the alternative way given on the right of Fig. 5. We have inserted in each edge of $\rho\tau_{3,2}$ an 'electric resistor'. There is also a resistor between the root and cross that highlights the root. Resistors in the graph correspond to factors ξ in (4.2), vertices in the graph connected to k resistors correspond to a k-th derivative of H. There is an index at each end of each resistors. Expressions like (4.2) will be called *Hamiltonian elementary differentials.* When the corresponding tree is of order r, the Hamiltonian elementary differential is said to be of order r.

4.2. Hamiltonian elementary Jacobians

Now assume that B-series (1.2) (with **f** = $\Xi \nabla H$) has been written in terms of the Hamiltonian elementary differentials; in order to check the symplecticness condition (1.6) we have to begin by computing the Jacobian matrix of the transformation defined by the B-series. Hence we have to find the Jacobian matrix of each Hamiltonian elementary differential.

Let us do this for (4.2). The component (i, z) in the Jacobian of this Hamiltonian elementary differential is obviously given by

$$\begin{aligned} & \xi_{iI} H_{Ijkz} \xi_{jJ} H_J \xi_{kK} H_K \\ + \quad & \xi_{iI} H_{Ijk} \xi_{jJ} H_{Jz} \xi_{kK} H_K \\ + \quad & \xi_{iI} H_{Ijk} \xi_{jJ} H_J \xi_{kK} H_{Kz} . \end{aligned} \tag{4.3}$$

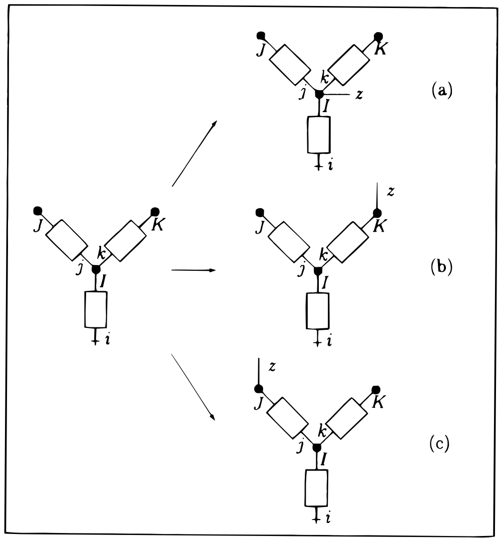

Fig. 6. Computation of the Jacobian of a Hamiltonian elementary differential

We may then visualize the process of computing the Jacobians as in Fig. 6. A 'branch' with the index z is successively born from each vertex. The index i attached to the free end of the root resistor and the index z at the new branch identify the components of the Jacobian. There is summation in all the remaining indices.

Before we proceed, it is important to realize that in (4.3) the second and third rows are equal; they only differ in the notation for the summation indices (the roles of (j, J) and (k, K) are interchanged). Of course this is because the graphs (b) and (c) in Fig. 6 are essentially the same; they only differ in the way they have been depicted. We then rewrite (4.3) as

$$\xi_{iI} H_{Ijkz} \xi_{jJ} H_J \xi_{kK} H_K + 2\xi_{iI} H_{Ijk} \xi_{jJ} H_{Jz} \xi_{kK} H_K =: D^1_{iz} + 2D^2_{iz}. \tag{4.4}$$

Each of the matrices D^1 and D^2 are called *(Hamiltonian) elementary Jacobians.* The Jacobian matrix of a Hamiltonian elementary differential is therefore a linear combination with integer coefficients of (distinct) Hamiltonian elementary Jacobians. The order of an elementary Jacobian is the order of the corresponding elementary differential. Thus the elementary Jacobians in (4.4) are of order 3.

Going back to (4.3), the reason why the last two rows are equal can be restated by saying that there is a *symmetry* of the rooted tree $\rho\tau_{3,2}$ that maps one of the end vertices into the other. In other words the end vertices belong to the same equivalence class of vertices, as defined in the previous section. When differentiating the Hamiltonian elementary differential associated with the rooted tree $\rho\tau$, the number of resulting distinct elementary Jacobians equals the number of equivalence classes in $\rho\tau$. Each (distinct) elementary Jacobian appears $S(v, \rho\tau)$ times, if v is the vertex where the branch indicating differentiation is appended.

To sum up, the (i, z) component of the Jacobian of (1.2) is

$$\Psi'_{iz} = \delta_{iz} + \sum_{n=1}^{\infty} \frac{h^n}{n!} \sum_{\rho\tau \in RT_n} c(\rho\tau) \sum S(v, \rho\tau) D^{(v,\rho\tau)}_{iz}. \tag{4.5}$$

Here δ_{iz} denotes the Kronecker symbol, the inner summation is extended to the classes of equivalent vertices v in $\rho\tau$ and $D_{iz}^{(v,\rho\tau)}$ denotes the (i,z) component of the elementary Jacobian obtained by differentiating at vertex v the elementary differential associated with $\rho\tau$.

In view of (4.5), it is useful to regard the identity matrix (δ_{iz}) as the unique elementary Jacobian of order 0 and set $D_{iz}^{\emptyset} = \delta_{iz}$.

4.3. Elementary products

In order to check the condition (1.6) we have to see whether for each $z^*, z = 1, \dots, 2d$

$$\psi'_{i^*z^*}\xi_{i^*i}\psi'_{iz} = \xi_{z^*z}, \tag{4.6}$$

with ψ'_{iz} given by (4.5). It is then useful to begin by computing quantities like

$$D_{i^*z^*}^{(v^*,\rho\tau^*)}\xi_{i^*i}D_{iz}^{(v,\rho\tau)}$$

for pairs of elementary Jacobians $D^{(v^*,\rho\tau^*)}$, $D^{(v,\rho\tau)}$. We first do this in some examples.

For the elementary Jacobians D^1 and D^2 in (4.4) we find

$$D^1_{i^*z^*}\xi_{i^*i}D^2_{iz} = (\xi_{i^*I^*}H_{I^*j^*k^*z^*}\xi_{j^*J^*}H_{J^*}\xi_{k^*K^*}H_{K^*})\xi_{i^*i}(\xi_{iI}H_{Ijk}\xi_{jJ}H_{Jz}\xi_{kK}H_K).$$

We may use the identity $\xi_{i^*i}\xi_{i^*I^*} = \delta_{iI^*}$ to get some simplification: After some rearrangement, we obtain

$$D^1_{i^*z^*}\xi_{i^*i}D^2_{iz} = H_{z^*j^*k^*i}\xi_{j^*J^*}H_{J^*}\xi_{k^*K^*}H_{K^*}\xi_{iI}H_{Ijk}\xi_{jJ}H_{Jz}\xi_{kK}H_K \tag{4.7}$$

The process is illustrated in Fig. 7(b). Two resistors have merged and disappeared.

The expression on the right of (4.7) can be associated with the 'tree with two branches' in the bottom right corner of Fig. 7. Between adjacent vertices there is a resistor, with an index at each end. There are two branches labelled z^* and z. In the factors ξ_{ab} corresponding to the resistors, a is the index 'closest' to the branch z^*. Vertices correspond to derivatives of H.

As a second example, we find (Fig. 7(a)):

$$\begin{aligned} D^2_{i^*z^*}\xi_{i^*i}D^2_{iz} &= (\xi_{i^*I^*}H_{I^*j^*k^*}\xi_{j^*J^*}H_{J^*z^*}\xi_{k^*K^*}H_{K^*})\xi_{i^*i}(\xi_{iI}H_{Ijk}\xi_{jJ}H_{Jz}\xi_{kK}H_K) \\ &= H_{z^*J^*}\xi_{j^*J^*}H_{j^*ik^*}\xi_{k^*K^*}H_{K^*}\xi_{iI}H_{Ijk}\xi_{jJ}H_{Jz}\xi_{kK}H_K \end{aligned}$$

In order to relate the last expression to the two-branch tree at the top right corner of Fig. 7, we replace $\xi_{j^*J^*}$ by $-\xi_{J^*j^*}$ so as to obey the rule that in ξ_{ab} the first index a is closest to the z^*. Thus

$$D^2_{i^*z^*}\xi_{i^*i}D^2_{iz} = -H_{z^*J^*}\xi_{J^*j^*}H_{j^*ik^*}\xi_{k^*K^*}H_{K^*}\xi_{iI}H_{Ijk}\xi_{jJ}H_{Jz}\xi_{kK}H_K). \tag{4.8}$$

The matrices in the right hand sides of (4.7) and (4.8) (without the sign) will be called elementary products of order 6. In general, ${D^{(v^*,\rho\tau^*)}}^{\mathrm{T}}\Xi D^{(v,\rho\tau)}$ equals plus or minus an elementary product of order equal to $n(\rho\tau^*) + n(\rho\tau)$. The sign is + if between the branch z^* and the root in $\rho\tau^*$ there is an even number of resistors (as in the bottom left corner of Fig. 7, where there are 0 resistors). The sign is $-$ for cases with an odd number of resistors (as in the top left corner of Fig. 7, with the (j^*, J^*) resistor).

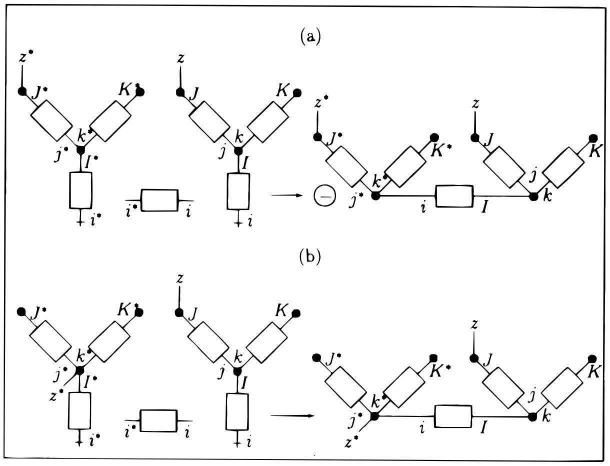

Fig. 7. Two instances of the computation of an elementary product

The construction of the elementary products also works in cases where either $\rho\tau^*$ or $\rho\tau$ (but not both) are the empty rooted tree with elementary Jacobian $D^\emptyset$.

Different combinations of elementary Jacobians may give rise to the same elementary product. For instance in Fig. 8 we see that there are 4 combinations that lead to the elementary product ${D^1}^{\mathrm{T}}\Xi D^2$ we have considered before. For each combination we have indicated whether the product arises with a + or with a $-$ sign.

In general, in the situation of Fig. 4 we have

$$\begin{aligned} -{D^\emptyset}^{\mathrm{T}}\Xi D^{(i_k,\rho\tau_{i_1})} &= {D^{(i_1,\rho\tau_{I_1-})}}^{\mathrm{T}}\Xi D^{(i_k,\rho\tau_{I_2+})} \\ &= -{D^{(i_1,\rho\tau_{I_2-})}}^{\mathrm{T}}\Xi D^{(i_k,\rho\tau_{I_3+})} \\ &= \dots\dots\dots \\ &= (-1)^{k+1}{D^{(i_1,\rho\tau_{i_k})}}^{\mathrm{T}}\Xi D^\emptyset. \end{aligned} \tag{4.9}$$

To prove the sufficiency part of Theorem 2.1, we replace in (4.6), $\psi'_{i^*z^*}$, ψ'_{iz} by their expansions in powers of h as in (4.5), and multiply termwise the series. In the result, there is a unique term containing the power h^0; this is $\delta_{i^*z^*}\xi_{i^*i}\delta_{iz} = \xi_{z^*z}$ and just matches the right hand side of (4.6). In the terms containing the power h^m, $m \geq 1$, we group together the contributions to the same elementary product of order m. By (4.9), the coefficient of each elementary product is a sum of the form considered in the left hand side of the identity in Lemma 3.5. This lemma shows that, under the condition of the main theorem, all those sums vanish. Hence (4.6) holds and the B-series is canonical.

5. Proof of the main result: Necessity

Clearly, the necessity of the condition in Theorem 2.1 is a consequence of the following result, that states the independence of distinct elementary products.

Lemma 5.1. *Let Π be an elementary product of order k, ($k = 1, 2, \ldots$). Then there is a (polynomial) Hamiltonian function H with $d = k + 1$ degrees of freedom for which*

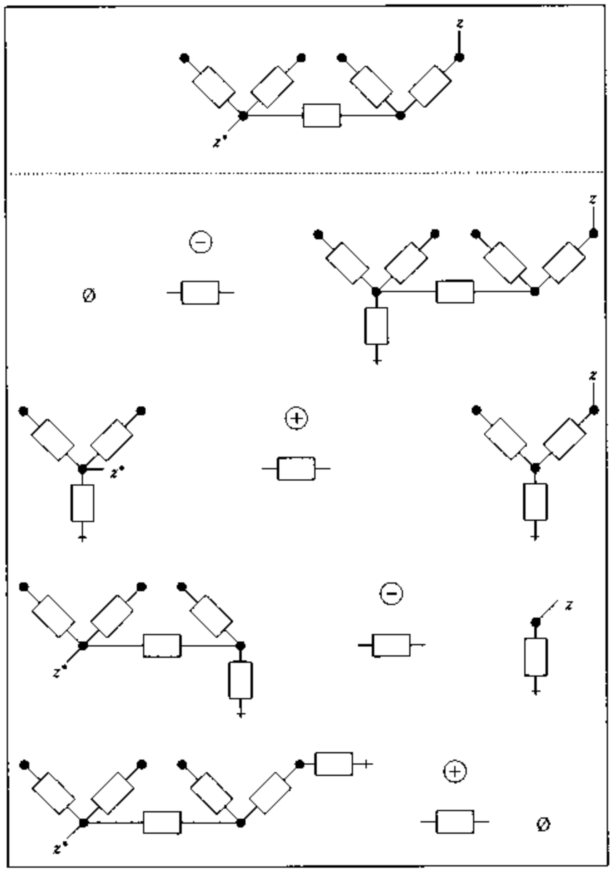

Fig. 8. Shown below the dotted line are 4 different combinations of elementary Jacobians leading to the same elementary product. The two-branch tree associated with the elementary product is displayed above the dotted line

the entry (1,2) of the matrix Π *at* $\mathbf{y} = \mathbf{0}$ *is* $\neq 0$*, while all other elementary products of order* $k+1$ *have vanishing (1,2) entry at* $\mathbf{y} = \mathbf{0}$.

Proof. Even though the proof is completely general, the underlying idea is best presented in an example. Let us consider the two-branch tree in Fig. 9, where $k = 5$. There are $k-1$ resistors in the graph. Each resistor has an 'entry' end, i.e., an end closest to the vertex with the z^* branch. At each of the $k-1$ entry ends we attach the numbers $3, 4, \dots, k+1$. To the 'exit' end of the resistor whose entry has been labelled i, we attach the number $i+d = i+k+1$. Finally attach the label 1 to the z^* branch and the label 2 to the z branch.

We now form the Hamiltonian

$$H = y^1 y^3 y^4 + y^{10} + y^5 y^6 y^9 + y^{11} + y^{12} y^2.$$

There are as many terms being summed as vertices in the graph; a vertex connected to branches or resistor ends $i_1, i_2, \dots, i_l$ introduces a term $y^{i_1} \cdots y^{i_l}$.

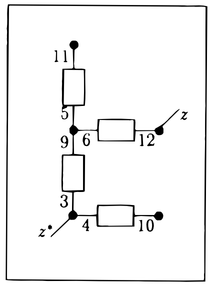

Fig. 9. Graph for the proof of necessity

Assume that, for this Hamiltonian and at $\mathbf{y} = \mathbf{0}$, an elementary product Π^* of order $k+1$ has nonzero (1, 2) entry. We show that the two-branch tree for Π^* is the same as the graph for Π and hence $\Pi^* = \Pi$.

By construction of H, derivatives of H of the form $H_{1,i_1,\cdots,i_m}$ are 0 at the origin, except for H_{134}. This shows that the vertex of the z^* branch is linked in the graph to two vertices. Furthermore

$$\Pi^*_{12} = H_{134}\xi_{3I}H_{I\ldots}\cdots\xi_{4J}H_{J\ldots} + H_{143}\xi_{4I}H_{I\ldots}\cdots\xi_{3J}H_{J\ldots}$$

At the origin, one of the terms in the right hand side is nonzero. Without loss of generality we may assume that the first is nonzero. (If the second is nonzero we change the names of the summation indices.) Hence

$$H_{134}\xi_{3I}H_{I\ldots}\cdots\xi_{4J}H_{J\ldots} \neq 0 \tag{5.1}$$

and in view of the structure of Ξ in (1.5), all terms in the sum (5.1) vanish except the term with $I = 3 + d = 9$. In turn, all derivatives of the form $H_{9,i_1,\cdots,i_m}$ at $\mathbf{y} = \mathbf{0}$ vanish except H_{956}. By now, we have proved that in the two-branch tree for π^*, the z^*-branch vertex is joined to two other vertices and one of these is also adjacent to three vertices. The iteration of this argument concludes the proof.

6. Application to Runge-Kutta methods

In this section, we assume that the B-series (1.2) is generated by an s-stage RK method with weights b_i and coefficient matrix (a_{ij}). Then it is well known that the quantities $c(\rho\tau)/\alpha(\rho\tau)\gamma(\rho\tau)$ in (2.1) are simply the elementary weights $\Phi(\rho\tau)$. Hence (2.1) reads

$$\Phi(\rho\tau_i) + \Phi(\rho\tau_j) = \Phi(\rho\tau_I)\Phi(\rho\tau_J). \tag{6.1}$$

Let us further, set

$$\rho\tau_I = [\rho\tau_1, \ldots, \rho\tau_m]$$

for suitable (not necessarily distinct and possibly empty) $\rho\tau_\alpha$, $\alpha = 1, \ldots, m$ and similarly

$$\rho\tau_J = [\rho\tau^1, \ldots, \rho\tau^n].$$

Then

$$\begin{aligned} \rho\tau_i &= [\rho\tau_J, \rho\tau_1, \ldots, \rho\tau_m], \\ \rho\tau_j &= [\rho\tau_I, \rho\tau^1, \ldots, \rho\tau^n] \end{aligned}$$

and, according to the definition of elementary weight, (6.1) becomes

$$\sum_{ij} b_i a_{ij} \prod_\alpha \Phi_i(\rho\tau_\alpha) \prod_\beta \Phi_j(\rho\tau^\beta) + \sum_{ij} b_j a_{ji} \prod_\alpha \Phi_i(\rho\tau_\alpha) \prod_\beta \Phi_j(\rho\tau^\beta) = \\ (\sum_i b_i \prod_\alpha \Phi_i(\rho\tau_\alpha))(\sum_j b_j \prod_\beta \Phi_j(\rho\tau^\beta))$$

where Φ_i denotes the elementary weight of the i-th stage. We rearrange as follows

$$\sum_{ij}(\prod_\alpha \Phi_i(\rho\tau_\alpha))(b_i a_{ij} + b_j a_{ji} - b_i b_j)(\prod_\beta \Phi_j(\rho\tau^\beta)) = 0. \tag{6.2}$$

(If $\rho\tau_I = [\emptyset]$, then $\prod_\alpha \Phi_i(\rho\tau_\alpha)$ is of course 1, and similarly for $\rho\tau_J$.)

From (6.2) we conclude that the condition

$$m_{ij} := b_i a_{ij} + b_j a_{ji} - b_i b_j = 0, \qquad 1 \leq i, j \leq s \tag{6.3}$$

implies the canonicity of the B-series. Of course (6.3) is the well-known sufficient condition for an RK method to be symplectic discovered by Lasagni [5], Sanz-Serna [6] and Suris [9]] It is known that this condition is not necessary in the uninteresting case where the RK method contains redundant stages. For methods with independent stages Lasagni (see [1]) showed that (6.3) is also necessary for canonicity. The available proof of necessity is rather lengthy and delicate.

Let us now provide an alternative necessity proof by using the theory of B-series. Assume then that a RK method without equivalent stages ([2] Theorem 383B) is canonical and consider the subspace X of $\mathbb{R}^s$ spanned by the vectors with components

$$\prod_{\rho\tau \in \bar{RT}} \Phi_i(\rho\tau), \qquad i = 1, \ldots, s,$$

where $\bar{RT}$ is any finite collection of rooted trees. The case where $\bar{RT}$ is empty is allowed with the standard convention that

$$\prod_{\rho\tau \in \emptyset} \Phi_i(\rho\tau) = 1.$$

The subspace X contains the vector $\mathbf{e} = [1, 1, \ldots, 1]^{\mathrm{T}}$ and is obviously closed with respect to componentwise multiplication of vectors. Furthermore for any two indices $i, j = 1, 2, \ldots, s$, there is a vector $\mathbf{v} = [v^1, v^2, \ldots, v^s]^{\mathrm{T}}$ with $v^i \neq v^j$. In fact since the i-th stage is different from the j-th stage, $\Phi_i(\rho\tau) \neq \Phi_j(\rho\tau)$ for some rooted tree $\rho\tau$. The Stone-Weierstrass theorem ([2], Sect. 306) entails that $X = \mathbf{R}^s$. Then (6.2) is telling us that the bilinear form with matrix (m_{ij}) vanish in the whole of $\mathbb{R}^s$. From here $m_{ij} = 0$, $i, j = 1, \ldots, s$, and (6.3) follows.

Acknowledgement. This research has been supported by project PB89-0351 (Dirección General de Investigación Científica y Técnica).

References

1. Abia, L., Sanz-Serna, J.M. (1993) Partitioned Runge-Kutta methods for separable Hamiltonian problems. Math. Comput. **60**, 617–634
2. Butcher, J.C. (1987) The Numerical Analysis of Ordinary Differential Equations, John Wiley, Chichester
3. Hairer, E., Nørsett S.P., Wanner, G. (1987) Solving Ordinary Differential Equations I, Nonstiff Problems, Springer, Berlin
4. Hairer, E., Wanner, G. (1974) On the Butcher group and general multivalue methods. Computing **13**, 1–15
5. Lasagni, F.M. (1988) Canonical Runge-Kutta methods. ZAMP **39**, 952–953
6. Sanz-Serna, J.M. (1988) Runge-Kutta schemes for Hamiltonian problems. BIT **28**, 877–883
7. Sanz-Serna, J.M. (1992) Symplectic Integrators for Hamiltonian problems: an overview. Acta Numerica **1**, 243–286
8. Sanz-Serna, J.M., Abia, L. (1991) Order conditions for canonical Runge-Kutta schemes. SIAM J. Numer. Anal. **28**, 1081–1096
9. Suris, Y.B. (1989) The canonicity of mappings generated by Runge-Kutta type methods when integrating the system $\ddot{x} = \partial U/\partial x$. U.S.S.R. Comput. Maths. and Math. Phys. **29**, 138–144

This article was processed by the author using the LaTeX style file *pljour1* from Springer-Verlag.