BACKWARD ANALYSIS OF NUMERICAL INTEGRATORS AND SYMPLECTIC METHODS

ERNST HAIRER[†]

Dedicated to John Butcher on the occasion of his 60th birthday

Abstract. A backward analysis of integration methods, whose numerical solution is a P-series, is presented. Such methods include Runge-Kutta methods, partitioned Runge-Kutta methods and Nyström methods. It is shown that the numerical solution can formally be interpreted as the exact solution of a perturbed differential system whose right-hand side is again a P-series. The main result of this article is that for symplectic integrators applied to Hamiltonian systems the perturbed differential equation is a Hamiltonian system too. The proofs use the one-to-one correspondence between rooted trees and the expressions appearing in the Taylor expansions of the exact and numerical solutions (elementary differentials).

Key words. Backward analysis, Hamiltonian systems, Runge-Kutta methods, symplectic methods, P-series.

1. Introduction

For the numerical solution of ordinary differential equations

$$y' = f(y) \tag{1.1}$$

we consider one-step integrators such as Runge-Kutta methods or something similar. The aim of this article is to investigate the following question: *can the numerical solution be interpreted as the exact solution of a perturbed differential equation?* E.g., for the explicit Euler discretization

$$y_{n+1} = y_n + hf(y_n) \tag{1.2}$$

we shall see (Griffiths & Sanz-Serna [GS86]) that the values y_n approximate $\widetilde{y}(t_n)$ with an error of $\mathcal{O}(h^2)$, where $\widetilde{y}(t)$ is the exact solution of the perturbed differential equation (method of modified equations, see also [WH74])

$$\widetilde{y}\,' = f(\widetilde{y}) - \frac{h}{2} f'(\widetilde{y}) f(\widetilde{y}). \tag{1.3}$$

Addition of a suitable $\mathcal{O}(h^2)$ term on the right-hand side of (1.3) allows to reduce the difference $y_n - \widetilde{y}(t_n)$ to $\mathcal{O}(h^3)$, etc. Section 2 of this paper elaborates these ideas for general integration methods. Explicit formulas for the perturbed differential equations are given. In view of an application of this theory to Hamiltonian systems, these results are extended to partitioned ordinary differential equations in Section 3.

The main interest of this article is in the study of the numerical solution of *Hamiltonian systems*

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

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

(in applications to mechanics, the variables $q=(q_1,\dots,q_n)^T$ are the generalized coordinates and $p=(p_1,\dots,p_n)^T$ are the generalized momenta). The flow of the system (1.4) has some very interesting properties: it is symplectic (i.e., in a certain fashion area-preserving; see Arnol'd [Ar89]) and preserves the Hamiltonian function, i.e., $H\big(p(t),q(t)\big)=Const$ along a trajectory of (1.4). In the last years several authors have investigated numerical methods which preserve the symplectic structure of the flow of (1.4). Characterizations for Runge-Kutta methods, for partitioned Runge-Kutta methods and for Nyström methods have been obtained by Sanz-Serna [SS88], Lasagni [La88], Suris [Su89]; a survey on these results is given in [SS92], see also [HNW93].

But what is known about the preservation of the Hamiltonian? Consider for example the system (1.4) with Hamiltonian function

$$H(p,q)=\frac{p^2}{2\,(1+U'(q)^2)}+U(q),\qquad U(q)=0.1\cdot\big(q\cdot(q-2)\big)^2+0.008\cdot q^3. \tag{1.5}$$

It describes the movement of a bead on a wire with form $y=U(q)$ in a gravitational field ([Ar89], page 102). To this problem we apply the well-known explicit and implicit Euler methods as well as the symplectic Euler method defined by

$$p_{n+1}=p_n-h\frac{\partial H}{\partial q}(p_{n+1},q_n),\qquad q_{n+1}=q_n+h\frac{\partial H}{\partial p}(p_{n+1},q_n). \tag{1.6}$$

We use the initial values $p_0=0.49$, $q_0=0$, a constant step size $h=1/4$ and integrate over the interval $[0,125]$. The numerical Hamiltonian $H(p_n,q_n)$ is plotted in Fig. 1. Whereas the values for the explicit and implicit Euler methods drift away from the exact value $H(p_0,q_0)$, the error $H(p_n,q_n)-H(p_0,q_0)$ remains nicely bounded for the symplectic method, also for very long integration intervals.

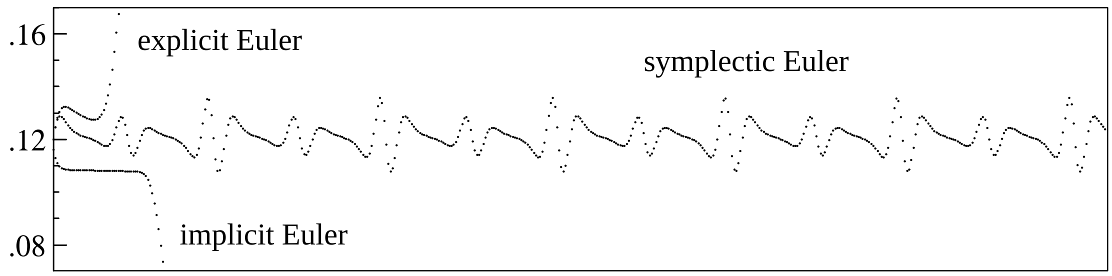


Fig. 1. Numerical Hamiltonian ($H(p_0,q_0)=0.12005$)

This phenomenon can be explained by a backward analysis argument ([Yo93], [SS92], [MK92], [AF91]) as follows: whenever we apply a symplectic integration method to a Hamiltonian system (1.4), then the numerical solution is close to the exact solution of a perturbed Hamiltonian system. E.g., for the symplectic Euler method (1.6), the perturbed Hamiltonian is given by

$$\widetilde{H}=H-\frac{h}{2}H_pH_q+\frac{h^2}{12}\Big(H_{pp}H_q^2+H_{qq}H_p^2+4H_{pq}H_qH_p\Big)+\dots\,. \tag{1.7}$$

The left picture of Fig. 2 shows the level curves of the Hamiltonian (1.5). To the right we have plotted the level curves of (1.7) with $h=0.4$ and we have included the numerical solution of (1.6) applied to (1.4) with constant step size $h=0.4$ and initial

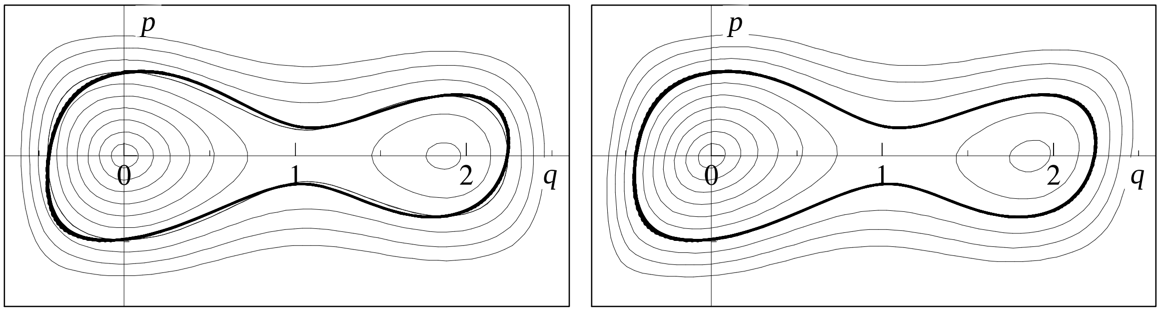


Fig. 2. Level curves of $H(p,q)$ (left) and $\widetilde{H}(p,q)$ (right) together with numerical solution

values $p_0 = 0.49$, $q_0 = 0$. An excellent agreement with the above statement can be observed.

The main result of this paper (Theorem 9, Section 5) states that the above backward analysis argument is possible for every symplectic integration method applied to an arbitrary Hamiltonian system. Explicit formulas for $\widetilde{H}(p,q)$ will be given. For the proof of this result we use a characterization of symplectic methods in terms of the coefficients of the Taylor series (P-series) of the numerical solution (Section 4). Technical details are postponed to Section 6. The proofs are based on a one-to-one correspondence between the expressions in the Taylor series of the exact and numerical solutions (elementary differentials) and rooted trees ([Bu87], [Ha81], [HNW93]).

2. Backward Analysis of Numerical Integrators

Let us first illustrate the techniques of this section at the explicit Euler discretization (1.2). The idea is to search for an h-dependent vector field $f_h(y)$ of the form

$$f_h(y) = \beta_1 f(y) + \beta_2 \frac{h}{2} f'(y)f(y) + \mathcal{O}(h^2)$$

and to determine the parameters β_1, β_2 in such a way that the numerical solution agrees with the exact solution $\widetilde{y}(x_0+h)$ of $\widetilde{y}\,' = f_h(\widetilde{y})$ as far as possible. Comparing the Taylor series of $\widetilde{y}(x_0+h)$

$$\begin{aligned}\widetilde{y}(x_0+h) &= y_0 + hf_h(y_0) + \frac{h^2}{2!} f_h'(y_0)f_h(y_0) + \mathcal{O}(h^3)\\ &= y_0 + h\beta_1 f(y_0) + \frac{h^2}{2!}(\beta_1^2+\beta_2)f'(y_0)f(y_0) + \mathcal{O}(h^3)\end{aligned}$$

with the numerical solution $y_1 = y_0 + hf(y_0)$ yields $\beta_1 = 1$, $\beta_2 = -1$ and formula (1.3) is obtained.

This idea can be extended to more general methods, e.g., Runge-Kutta methods. For the problem $y' = f(y)$ they are defined by

$$y_1 = y_0 + h\sum_{i=1}^{s} b_i k_i, \qquad k_i = f\Big(y_0 + h\sum_{j=1}^{s} a_{ij}k_j\Big). \tag{2.1}$$

Since the fundamental work of Butcher in the early sixties it is known that the Taylor series of the numerical solution y_1 (considered as a function of the step size h) can be written as

$$y_1 = y_0 + \sum_{t\in T} \frac{h^{\rho(t)}}{\rho(t)!}\,\alpha(t)\Big(\gamma(t)\sum_{i=1}^{s} b_i\Phi_i(t)\Big)F(t)(y_0). \tag{2.2}$$

The symbols appearing in this formula have the following meaning:

- T set of rooted trees;
- $\rho(t)$ number of vertices of the tree t (order of t);
- $\alpha(t)$ number of possible monotonic labellings of t;
- $\gamma(t)$ an integer coefficient;
- $\Phi_i(t)$ an expression depending on the coefficients of the method;
- $F(t)$ the elementary differential, depending on the function $f(y)$.

It is common to denote by τ the only tree of order 1, and by

$$t = [t_1,\ldots,t_m] \tag{2.3}$$

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. Using this representation of trees, the above expressions can be recursively defined as follows: for the tree τ we have

$$\rho(\tau)=1,\qquad \alpha(\tau)=1,\qquad \gamma(\tau)=1,\qquad \Phi_i(\tau)=1,\qquad F(\tau)(y)=f(y)$$

and for $t=[t_1,\ldots,t_m]$ it holds

$$\rho(t) = 1+\rho(t_1)+\ldots+\rho(t_m), \tag{2.4}$$

$$\alpha(t) = \binom{\rho(t)-1}{\rho(t_1),\ldots,\rho(t_m)}\cdot\alpha(t_1)\cdot\ldots\cdot\alpha(t_m)\cdot\frac{1}{\mu_1!\mu_2!\ldots}, \tag{2.5}$$

$$\gamma(t) = \rho(t)\cdot\gamma(t_1)\cdot\ldots\cdot\gamma(t_m), \tag{2.6}$$

$$\Phi_i(t) = \sum_{j_1,\ldots,j_m} a_{ij_1}\Phi_{j_1}(t_1)\cdot\ldots\cdot a_{ij_m}\Phi_{j_m}(t_m), \tag{2.7}$$

$$F(t)(y) = f^{(m)}(y)\cdot\Big(F(t_1)(y),\ldots,F(t_m)(y)\Big). \tag{2.8}$$

The integers $\mu_1,\mu_2,\ldots$ in (2.5) count the numbers of equal trees among $t_1,\ldots,t_m$. For more details on these definitions and for a proof of formula (2.2) we refer the reader to Butcher [Bu87] or to the monograph of Hairer, Nørsett & Wanner [HNW93], the notation of which we shall use throughout this paper.

The results of this section are not restricted to Runge-Kutta methods. All what we need is that the numerical solution after one step is given by a formal series

$$y_1 = y_0 + \sum_{t\in T}\frac{h^{\rho(t)}}{\rho(t)!}\,\alpha(t)\,a(t)\,F(t)(y_0) \tag{2.9}$$

where $a: T\to\mathbb{R}$ is an arbitrary mapping. Such a series is called a B-series. If the function $f(y)$ is only N times continuously differentiable, then the series (2.9) has to be interpreted as a finite sum (over $t\in T$ with $\rho(t)\le N$) with a remainder of size $\mathcal{O}(h^{N+1})$.

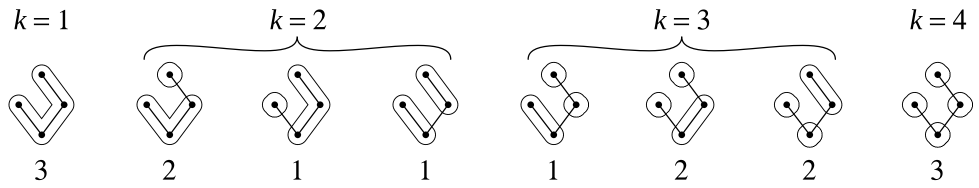


Fig. 3. All partitions of $t=[\tau,[\tau]]$ into k subtrees together with the numbers $\alpha(t,S)$

Definition. Let $t\in T$ be a rooted tree. A *partition* of t into k subtrees $s_1,\dots,s_k$ is a set S consisting of $k-1$ branches of t such that the trees $s_1,\dots,s_k$ are obtained when the branches of S are removed from t. Such a partition is denoted by (t,S). We further denote by $\alpha(t,S)$ the number of possible monotonic labellings of t such that the vertices of the subtrees s_j are labelled consecutively.

Example. Consider the tree $t=[\tau,[\tau]]$ of order $\rho(t)=4$ satisfying $\alpha(t)=3$. All possible partitions of t into k subtrees are shown in Fig. 3. The numbers $\alpha(t,S)$ are given below the corresponding trees. We observe that different partitions of t may lead to the same set of subtrees $\{s_1,\dots,s_k\}$.

Theorem 1. *Let $a:T\to\mathbb{R}$ be given and assume that the right-hand side of (1.1) is N times continuously differentiable. The numerical solution y_1, given by (2.9), then satisfies*

$$y_1=\widetilde{y}(x_0+h)+\mathcal{O}(h^{N+1}) \tag{2.10}$$

where $\widetilde{y}(x)$ is the exact solution of the perturbed differential equation

$$\widetilde{y}\,'=\sum_{\rho(t)\le N}\frac{h^{\rho(t)-1}}{\rho(t)!}\,\alpha(t)\,b(t)\,F(t)(\widetilde{y}). \tag{2.11}$$

The coefficients $b(t)$ are recursively defined by

$$a(t)=\sum_{k=1}^{\rho(t)}\frac{1}{k!}\sum_{(t,S)}\binom{\rho(t)}{\rho(s_1),\dots,\rho(s_k)}\frac{\alpha(t,S)}{\alpha(t)}\,b(s_1)\cdot\ldots\cdot b(s_k). \tag{2.12}$$

The second sum in (2.12) is over all partitions of t into k subtrees $s_1,\dots,s_k$.

Example. For the tree $t=[\tau,[\tau]]$ of Fig. 3 formula (2.12) becomes

$$a([\tau,[\tau]])=b([\tau,[\tau]])+\frac{1}{2!}\left(\binom{4}{3}\frac{2}{3}\,b([\tau,\tau])\,b(\tau)+\binom{4}{3}\frac{1}{3}\,b([[\tau]])\,b(\tau)+\binom{4}{2}\frac{1}{3}\,b([\tau])^2\right)$$
$$+\frac{1}{3!}\binom{4}{2,1,1}\frac{5}{3}\,b([\tau])\,b(\tau)^2+b(\tau)^4\,.$$

In the general situation we have the formula

$$a(t)=b(t)+\text{ expressions involving } b(s) \text{ with } \rho(s)<\rho(t) \tag{2.13}$$

which is displayed in Table 1 for trees up to order 3. Consequently, the coefficients

$b(t)$ are uniquely determined and can be computed recursively from (2.12).

Table 1. Examples for formula (2.12)

$t = \bullet$	$a(\bullet) = b(\bullet)$
$t = \diagup$	$a(\diagup) = b(\diagup) + b(\bullet)^2$
$t = \vee$	$a(\vee) = b(\vee) + \frac{3}{2} b(\diagup) b(\bullet) + b(\bullet)^3$
$t = \rangle$	$a(\rangle) = b(\rangle) + 3b(\diagup) b(\bullet) + b(\bullet)^3$

Remarks. a) For the explicit Euler method (1.2) the coefficients of the B-series (2.9) satisfy $a(\tau) = 1$ and $a(t) = 0$ for $\rho(t) \geq 2$. Inserting these values into the formulas of Table 1 we obtain

$$b(\bullet) = 1, \qquad b(\diagup) = -1, \qquad b(\vee) = 1/2, \qquad b(\rangle) = 2, \qquad \ldots$$

and the perturbed differential equation (2.11) becomes for $N = 3$ (compare (1.3))

$$\widetilde{y}\,' = f(\widetilde{y}) - \frac{h}{2} f'(\widetilde{y}) f(\widetilde{y}) + \frac{h^2}{12}\Big(\big(f''(f,f)\big)(\widetilde{y}) + 4\,(f'f'f)(\widetilde{y}) \Big).$$

b) For the implicit Euler method the coefficients of (2.9) satisfy $a^*(t) = \gamma(t)$ for all trees t. Computing the corresponding coefficients of the series (2.11) one observes that

$$b^*(t) = (-1)^{\rho(t)-1} b(t). \tag{2.14}$$

Relation (2.14) is true in general if $b(t)$ and $b^*(t)$ are the coefficients which correspond to adjoint methods (see [HNW93, Section II.8]). In particular, for symmetric methods (where the adjoint method coincides with the original one) it holds $b(t) = 0$ if $\rho(t)$ is even. Therefore the right-hand side of (2.11) contains only even powers of h in this situation.

c) If the method is of order p, i.e., $a(t) = 1$ for $\rho(t) \leq p$, then we have $b(\tau) = 1$, $b(t) = 0$ for $2 \leq \rho(t) \leq p$ and $b(t) = a(t) - 1$ for $\rho(t) = p + 1$. Hence, the perturbation in formula (2.11) is of size $\mathcal{O}(h^p)$ and the leading term is the same as for the local truncation error.

d) Using a standard convergence argument for one-step methods it follows from (2.10) that

$$y_n = \widetilde{y}(x_n) + \mathcal{O}(h^N)$$

for $x_n = x_0 + nh \leq Const$. This yields the asymptotic expansion of y_n because the solution $\widetilde{y}(x)$ of (2.11) can be expanded into an h-series (use for example the Gröbner-Alekseev formula [HNW93, Section I.14]).

e) The famous Euler-Maclaurin formula

$$\begin{aligned} T(h) &= \int_0^1 f(x)\,dx + \frac{h^2}{12}\Big(f'(1) - f'(0)\Big) - \frac{h^4}{720}\Big(f'''(1) - f'''(0)\Big) + \mathcal{O}(h^6) \\ &= \int_0^1 \Big(f(x) + \frac{h^2}{12} f''(x) - \frac{h^4}{720} f^{(4)}(x)\Big)\,dx + \mathcal{O}(h^6) = \int_0^1 f_h(x)\,dx + \mathcal{O}(h^6) \end{aligned}$$

represents the numerical result of the trapezoidal rule as the exact integral over a perturbed function $f_h(x)$. It can be considered as a special case of Theorem 1.

Indeed, for the problem $y' = f(x)$, $y(0) = 0$ all elementary differentials vanish with the exception of those for the bushy trees $\tau_1 = \tau$, $\tau_2 = [\tau]$, $\tau_3 = [\tau, \tau]$, $\tau_4 = [\tau, \tau, \tau]$, etc. It holds $F(\tau_k)(y) = f^{(k-1)}(x)$, so that the integration of the series (2.11) leads to

$$\widetilde{y}(x) = b(\tau)\int_0^x f(\xi)\,d\xi + \sum_{k=2}^{N} \frac{h^{k-1}}{k!}\, b(\tau_k)\Big(f^{(k-2)}(x) - f^{(k-2)}(0)\Big) + \mathcal{O}(h^N)$$

which, for the trapezoidal rule, becomes the above formula. Consequently, the coefficients $b(\tau_k)$ are related to the Bernoulli numbers and the infinite series

$$\sum_{t\in T} \frac{h^{\rho(t)-1}}{\rho(t)!}\, \alpha(t)\, b(t)\, F(t)(\widetilde{y}) \tag{2.15}$$

does not converge in general. This is the reason why we have to work with truncated series in Theorem 1.

Proof of Theorem 1. Assuming sufficient differentiabilty of $f(y)$ we shall prove that the derivatives of $\widetilde{y}(x)$ satisfy

$$\frac{h^k}{k!}\,\widetilde{y}^{\,(k)}(x) = \sum_{t\in LT, k\le\rho(t)\le N} \frac{h^{\rho(t)}}{\rho(t)!}\, b^{(k)}(t)\, F(t)\big(\widetilde{y}(x)\big) + \mathcal{O}(h^{N+1}) \tag{2.16}$$

with coefficients $b^{(k)}(t)$ to be constructed below (in this part of the proof we work with labelled trees $t \in LT$ and therefore omit the factor $\alpha(t)$). By definition of $\widetilde{y}(x)$ this formula is true for $k=1$ with $b^{(1)}(t) = b(t)$. For general k we differentiate formula (2.11) to obtain

$$\frac{h^k}{k!}\,\widetilde{y}^{\,(k)}(x) = \frac{1}{k!}\sum_{\rho(u)\le N} \frac{h^{\rho(u)}}{\rho(u)!}\, b(u)\cdot h^{k-1}\cdot \frac{d^{k-1}}{dx^{k-1}}\Big(F(u)\big(\widetilde{y}(x)\big)\Big). \tag{2.17}$$

Therefore we have to compute the derivatives of $F(u)\big(\widetilde{y}(x)\big)$. For a fixed labelled tree $u \in LT$ this expression consists of $\rho(u)$ factors. We differentiate it with the help of Leibniz' rule and the chain rule and at the same time we replace $h\widetilde{y}\,'(x)$ by formula (2.11). This yields

$$h\cdot\frac{d}{dx}\Big(F(u)\big(\widetilde{y}(x)\big)\Big) = \sum_{i=1}^{\rho(u)}\sum_{\rho(v)\le N} \frac{h^{\rho(v)}}{\rho(v)!}\, b(v)\, F(u\circ_i v)\big(\widetilde{y}(x)\big) \tag{2.18}$$

where $u\circ_i v$ denotes the labelled tree which is obtained by adding a new branch with v to the vertex of u with label i (the labels of v have to be increased by $\rho(u)$). We continue to differentiate (2.18) and each time we replace in the right-hand side the appearing first derivatives of $F(u)\big(\widetilde{y}(x)\big)$ by formula (2.18). In this way we obtain

$$\begin{aligned} h^l\cdot\frac{d^l}{dx^l}\Big(F(u_1)\big(\widetilde{y}(x)\big)\Big) = \sum_{u_2,\ldots,u_l\in LT}\sum_{i_1=1}^{\rho_1}\sum_{i_2=1}^{\rho_1+\rho_2}\cdots\sum_{i_l=1}^{\rho_1+\ldots+\rho_l} \frac{h^{\rho(u_2)+\ldots+\rho(u_l)}}{\rho(u_2)!\cdot\ldots\cdot\rho(u_l)!} \\ \cdot\, b(u_2)\cdot\ldots\cdot b(u_l)\cdot F(u_1\circ_{i_1} u_2\circ_{i_2}\ldots\circ_{i_{l-1}} u_l)\big(\widetilde{y}(x)\big) \end{aligned} \tag{2.19}$$

where the first sum is over those $u_2,\ldots,u_l\in LT$ for which $\rho_i = \rho(u_i)\le N$ and where we have used the notation $u\circ_i v\circ_j w := (u\circ_i v)\circ_j w$, etc. Inserting formula (2.19) with $l = k-1$ into (2.17) and collecting the coefficients of those tupels $(u_1, i_1, u_2, i_2, \ldots, u_l)$

for which $u_1 \circ_{i_1} u_2 \circ_{i_2} \ldots \circ_{i_{l-1}} u_l$, considered as a tree without labels, equals t, we get formula (2.16) with

$$b^{(k)}(t) = \frac{1}{k!} \sum_{(t,S)} \binom{\rho(t)}{\rho(s_1), \ldots, \rho(s_k)} \frac{\alpha(t,S)}{\alpha(t)}\, b(s_1) \cdot \ldots \cdot b(s_k).$$

As in (2.12) the sum is over all partitions of t into k subtrees $s_1, \ldots, s_k$.

Since by Taylor expansion

$$\widetilde{y}(x_0 + h) = y_0 + \sum_{k=1}^{N} \frac{h^k}{k!}\, \widetilde{y}^{(k)}(x_0) + \mathcal{O}(h^{N+1}),$$

we see from (2.16) and (2.9) that we obtain (2.10) by putting

$$a(t) = \sum_{k=1}^{\rho(t)} b^{(k)}(t).$$

But this is exactly formula (2.12).

If $f(y)$ is only N times continuously differentiable, then the right-hand side of (2.11) is once continuously differentiable and, although the solution $\widetilde{y}(x)$ exists in a neighbourhood of the initial value (x_0, y_0), it need not be more than twice differentiable. In this case we replace $f(y)$ by its truncated Taylor expansion (developed around y_0) and use the fact that the difference of the solutions at $x_0 + h$ is of size $\mathcal{O}(h^{N+1})$. □

3. Extension to Partitioned Systems

In this section we extend Theorem 1 to numerical methods for partitioned ordinary differential equations of the form

$$p' = f(p,q), \qquad q' = g(p,q). \tag{3.1}$$

In view of later applications to Hamiltonian systems, we denote the components of the solution by p and q, but we let for the moment the functions f and g be arbitrary.

For the numerical solution of (3.1) we consider partitioned Runge-Kutta methods, given by

$$\begin{aligned} p_1 &= p_0 + h \sum_{i=1}^{s} b_i k_i, \qquad & k_i &= f\Big(p_0 + h \sum_{j=1}^{s} a_{ij} k_j,\, q_0 + h \sum_{j=1}^{s} \widehat{a}_{ij} \ell_j\Big) \\ q_1 &= q_0 + h \sum_{i=1}^{s} \widehat{b}_i \ell_i, \qquad & \ell_i &= g\Big(p_0 + h \sum_{j=1}^{s} a_{ij} k_j,\, q_0 + h \sum_{j=1}^{s} \widehat{a}_{ij} \ell_j\Big). \end{aligned} \tag{3.2}$$

Expanding p_1 and q_1 into a Taylor series there appear expressions (elementary differentials) depending on f and on g. In order to represent them graphically we have to consider trees with two different kinds of vertices, one (meagre) representing f, the other (fat) representing g (see [HNW93, Section II.15]). We denote by TP the set of all such trees; by TP_p and TP_q those subsets which consist of trees with meagre and fat roots, respectively. We denote the trees with one vertex by $\tau_p = \bullet$ and $\tau_q = \circ$, and by

$$u = {}_p[t_1, \ldots, t_m] \qquad \text{and} \qquad v = {}_q[t_1, \ldots, t_m] \tag{3.3}$$

those trees which are obtained by connecting the roots of $t_1, \ldots, t_m \in TP$ with a new root which is meagre in the first case and fat in the second one.

The numerical solution of (3.2) can then be formally written as

$$\begin{aligned} p_1 &= p_0 + \sum_{u \in TP_p} \frac{h^{\rho(u)}}{\rho(u)!}\, \alpha(u)\, a(u) F(u)(p_0, q_0) \\ q_1 &= q_0 + \sum_{v \in TP_q} \frac{h^{\rho(v)}}{\rho(v)!}\, \alpha(v)\, a(v) F(v)(p_0, q_0) \end{aligned} \tag{3.4}$$

(use truncated series, if necessary) where the coefficients $a(u), a(v)$ satisfy

$$\begin{aligned} a(u) &= \gamma(u) \sum_{i=1}^{s} b_i \Phi_i(u) \qquad &&\text{for } u \in TP_p \\ a(v) &= \gamma(v) \sum_{i=1}^{s} \widehat{b}_i \Phi_i(v) \qquad &&\text{for } v \in TP_q. \end{aligned} \tag{3.5}$$

Series of the form (3.4) are called P-series. The numbers $\rho(t)$ and $\gamma(t)$ are exactly the same as in (2.2) and are independent of the type of the vertices. The coefficients $\alpha(t)$ again count how many monotonic labellings of t are possible. It also satisfies the recursion formula (2.5), but the numbers μ_i and hence also $\alpha(t)$ itself depends on the type of the vertices of t. E.g., it holds $\alpha({}_p[\tau_p,\tau_p]) = 1$ whereas $\alpha({}_p[\tau_q,\tau_p]) = 2$. The elementary differentials are defined by $F(\tau_p)(p,q) = f(p,q)$, $F(\tau_q)(p,q) = g(p,q)$ and

$$\begin{aligned} F(u)(p,q) &= \frac{\partial^{m+l} f(p,q)}{\partial p^m\, \partial q^l} \Big(F(u_1)(p,q), \ldots, F(v_1)(p,q), \ldots \Big) \\ &\qquad \text{if } u = {}_p[u_1, \ldots, u_m, v_1, \ldots, v_l],\ \ u_i \in TP_p,\ v_i \in TP_q, \\ F(v)(p,q) &= \frac{\partial^{m+l} g(p,q)}{\partial p^m\, \partial q^l} \Big(F(u_1)(p,q), \ldots, F(v_1)(p,q), \ldots \Big) \\ &\qquad \text{if } v = {}_q[u_1, \ldots, u_m, v_1, \ldots, v_l],\ \ u_i \in TP_p,\ v_i \in TP_q. \end{aligned} \tag{3.6}$$

Finally we have $\Phi_i(\tau_p) = \Phi_i(\tau_q) = 1$ and

$$\Phi_i(t) = \sum_{\substack{j_1, \ldots, j_m \\ k_1, \ldots, k_l}} a_{ij_1} \Phi_{j_1}(u_1) \ldots a_{ij_m} \Phi_{j_m}(u_m) \widehat{a}_{ik_1} \Phi_{k_1}(v_1) \ldots \widehat{a}_{ik_l} \Phi_{k_l}(v_l) \tag{3.7}$$

for $t = {}_p[u_1, \ldots, u_m, v_1, \ldots, v_l]$ or $t = {}_q[u_1, \ldots, u_m, v_1, \ldots, v_l]$ where $u_i \in TP_p$ and $v_i \in TP_q$. The expression $\Phi_i(t)$ does not depend on the type of the root.

A partition of $t \in TP$ into subtrees is defined in exactly the same manner as for $t \in T$ (see the definition before Theorem 1). Again we use the notation (t, S) for a partition and $\alpha(t, S)$ for the number of monotonic labellings of t such that the subtrees of the partition are labelled consecutively. The following extension of Theorem 1 considers general P-series (3.4). It is not assumed that the coefficients are given by (3.5).

Theorem 2. *Let $a : TP \to \mathbb{R}$ be given and assume that the right-hand sides of (3.1) are N times continuously differentiable. The numerical solution p_1, q_1, given by (3.4), then satisfies*

$$p_1 = \widetilde{p}(x_0 + h) + \mathcal{O}(h^{N+1}), \qquad q_1 = \widetilde{q}(x_0 + h) + \mathcal{O}(h^{N+1}), \tag{3.8}$$

where $\widetilde{p}(x), \widetilde{q}(x)$ is the exact solution of the perturbed differential equation

$$\begin{aligned} \widetilde{p}\,' &= \sum_{\rho(u) \le N} \frac{h^{\rho(u)-1}}{\rho(u)!} \, \alpha(u)\, b(u)\, F(u)(\widetilde{p}, \widetilde{q}), \\ \widetilde{q}\,' &= \sum_{\rho(v) \le N} \frac{h^{\rho(v)-1}}{\rho(v)!} \, \alpha(v)\, b(v)\, F(v)(\widetilde{p}, \widetilde{q}). \end{aligned} \tag{3.9}$$

The coefficients $b(t)$ for $t \in TP$ are given by the recursion formula (2.12). □

Table 2 presents the defining formulas for $b : TP \to \mathbb{R}$ for all trees $u \in TP_p$ of order $\rho(u) \le 3$. Those for $v \in TP_q$ have the same coefficients.

Table 2. Examples for the definition of $b : TP \to \mathbb{R}$

$t = \bullet$	$a(\bullet) = b(\bullet)$
$t = \bullet[\bullet]$	$a(\bullet[\bullet]) = b(\bullet[\bullet]) + b(\bullet)^2$
$t = \bullet[\circ]$	$a(\bullet[\circ]) = b(\bullet[\circ]) + b(\bullet)b(\circ)$
$t = \bullet[\bullet,\bullet]$	$a(\bullet[\bullet,\bullet]) = b(\bullet[\bullet,\bullet]) + \frac{3}{2}b(\bullet[\bullet])b(\bullet) + b(\bullet)^3$
$t = \bullet[\circ,\bullet]$	$a(\bullet[\circ,\bullet]) = b(\bullet[\circ,\bullet]) + \frac{3}{4}b(\bullet[\circ])b(\bullet) + \frac{3}{4}b(\bullet[\bullet])b(\circ) + b(\bullet)^2 b(\circ)$
$t = \bullet[\circ,\circ]$	$a(\bullet[\circ,\circ]) = b(\bullet[\circ,\circ]) + \frac{3}{2}b(\bullet[\circ])b(\circ) + b(\bullet)b(\circ)^2$
$t = \bullet[\bullet[\bullet]]$	$a(\bullet[\bullet[\bullet]]) = b(\bullet[\bullet[\bullet]]) + 3b(\bullet[\bullet])b(\bullet) + b(\bullet)^3$
$t = \bullet[\circ[\bullet]]$	$a(\bullet[\circ[\bullet]]) = b(\bullet[\circ[\bullet]]) + \frac{3}{2}b(\bullet[\circ])b(\bullet) + \frac{3}{2}b(\bullet)b(\circ[\bullet]) + b(\bullet)^2 b(\circ)$
$t = \bullet[\bullet[\circ]]$	$a(\bullet[\bullet[\circ]]) = b(\bullet[\bullet[\circ]]) + \frac{3}{2}b(\bullet[\bullet])b(\circ) + \frac{3}{2}b(\bullet)b(\bullet[\circ]) + b(\bullet)^2 b(\circ)$
$t = \bullet[\circ[\circ]]$	$a(\bullet[\circ[\circ]]) = b(\bullet[\circ[\circ]]) + \frac{3}{2}b(\bullet[\circ])b(\circ) + \frac{3}{2}b(\bullet)b(\circ[\circ]) + b(\bullet)b(\circ)^2$

4. Symplectic P-Series

This section is devoted to a characterization of symplectic partitioned Runge-Kutta methods in terms of the coefficients $a : TP \to \mathbb{R}$ of the P-series (3.4). We consider Hamiltonian systems written in the form (1.4).

Recall that a transformation $(p, q) \mapsto (\bar{p}, \bar{q})$ is symplectic if it preserves the differential 2-form

$$\omega^2 = \sum_{J=1}^{n} dp^J \wedge dq^J \tag{4.1}$$

(see e.g., [Ar89]). We have used superscripts for the components of p and q in order to distinguish them from the numerical solutions p_0, q_0, p_1, q_1.

Theorem 3. *Consider a partitioned Runge-Kutta method (3.2) applied to the Hamiltonian system (1.4). If the method satisfies*

$$b_i = \widehat{b}_i, \qquad i = 1, \dots, s \tag{4.2}$$

$$b_i \widehat{a}_{ij} + \widehat{b}_j a_{ji} - b_i \widehat{b}_j = 0, \qquad i, j = 1, \dots, s \tag{4.3}$$

then it is symplectic, i.e., the numerical flow $(p_0, q_0) \mapsto (p_1, q_1)$ *is symplectic.*

If the method is irreducible, then the conditions (4.2), (4.3) are also necessary for the symplecticity of the numerical flow. □

The sufficiency of conditions (4.2), (4.3) is due to Sanz-Serna [SS92] and Suris [Su90]; the necessity of (4.3) has been proved for separable Hamiltonians by Abia & Sanz-Serna [AS93]. Using the following definition of irreducible methods their proof can be straight-forwardly extended to the general situation.

Irreducible methods. Two stages i and j of the method (3.2) are said to be *equivalent* if for every initial value problem (3.1) and for every sufficiently small step size h it holds $k_i = k_j$ and $\ell_i = \ell_j$. A partitioned Runge-Kutta method (3.2) is called *irreducible* if it possesses no equivalent stages (see [Bu87, Section 383] for a detailed discussion of the irreducibility of standard Runge-Kutta methods). We shall need the following algebraic characterization of irreducibility.

Lemma 4. *Let* Φ *be the* $s \times \infty$ *matrix whose elements are* $\Phi_i(t)$ *where* $i = 1, \dots, s$ *and* $t \in TP_p$ *(or* $t \in TP_q$*). The partitioned Runge-Kutta method (3.2) is irreducible if and only if the matrix* Φ *has full rank* s*.*

Proof. If the stages i and j are equivalent, it follows from the expansion

$$k_i = \sum_{u \in TP_p} \frac{h^{\rho(u)}}{\rho(u)!} \, \alpha(u) \, \gamma(u) \, \Phi_i(u) \, F(u)(p_0 . q_0) \tag{4.4}$$

(see [HNW93, Section II.15]) and from the independency of the elementary differentials that $\Phi_i(u) = \Phi_j(u)$ for all $u \in TP_p$. Hence, the lines i and j of Φ are identical.

On the other hand, if the method is irreducible it holds: for every pair (i, j) with $i \neq j$ there exists a tree $u \in TP_p$ such that $\Phi_i(u) \neq \Phi_j(u)$. Consequently a certain finite linear combination of the columns of Φ has distinct elements, i.e., there exist vectors $\xi \in \mathbb{R}^\infty$ (only finitely many components are non zero) and $\eta \in \mathbb{R}^s$ with $\Phi\xi = \eta$ and $\eta_i \neq \eta_j$ for all $i \neq j$. Next we observe that the componentwise product of two columns of Φ is again a column of Φ. This implies that the vector $(\eta_1^2, \dots, \eta_s^2)^T$ can also be written as a finite linear combination of the columns of Φ. Continuing this argumentation and observing that $(1, \dots, 1)^T$ is the first column of Φ we obtain a matrix X such that

$$\Phi X = (\eta_i^{j-1})_{i,j=1}^s$$

is a Vandermonde matrix. Since the η_i are distinct, the matrix Φ must have full rank. □

Our next aim is to interpret the conditions of Theorem 3 in terms of the coefficients (3.5) of the P-series representing the numerical solution (p_1, q_1). Since neither $\gamma(t)$

nor $\Phi_i(t)$ depend on whether the root of t is meagre or fat, condition (4.2) implies that

$$a(t) \text{ is independent of the type of the root of } t.$$

We next multiply formula (4.3) by $\Phi_i(u)\cdot\Phi_j(v)$ (where $u = {}_p[u_1,\ldots,u_m] \in TP_p$ and $v = {}_q[v_1,\ldots,v_l] \in TP_q$) and sum over all i and j. Using the recursion (3.7) this yields

$$\sum_{i=1}^{s} b_i\Phi_i(u\circ v) + \sum_{j=1}^{s} \widehat{b}_j\Phi_j(v\circ u) - \Big(\sum_{i=1}^{s} b_i\Phi_i(u)\Big)\Big(\sum_{j=1}^{s} \widehat{b}_j\Phi_j(v)\Big) = 0. \tag{4.5}$$

Here we have used the notation of Butcher ([Bu87], Section 143)

$$u\circ v = {}_p[u_1,\ldots,u_m,v], \qquad v\circ u = {}_q[v_1,\ldots,v_l,u]. \tag{4.6}$$

With the help of (3.5), formula (4.5) can be written in terms of the coefficients $a(t)$. This yields the following characterization of symplectic P-series.

Theorem 5. *A partitioned Runge-Kutta method (3.2) is symplectic if and only if the coefficients of the P-series (3.4) satisfy*

$$a(t) \;\; \textit{is independent of the type of the root of } t, \tag{4.7}$$

$$\frac{a(u\circ v)}{\gamma(u\circ v)} + \frac{a(v\circ u)}{\gamma(v\circ u)} = \frac{a(u)}{\gamma(u)}\cdot\frac{a(v)}{\gamma(v)} \qquad \textit{for } \; u\in TP_p,\; v\in TP_q. \tag{4.8}$$

Proof. If the method is reducible, it is equivalent to one with fewer stages. The property of being symplectic as well as the coefficients $a: TP \to \mathbb{R}$ of the series (3.4) are independent of the representative within a class of equivalent methods. Therefore, we may assume without loss of generality that the method under consideration is irreducible.

We have shown above that the two conditions (4.2), (4.3) imply the properties (4.7), (4.8). Hence, an application of Theorem 3 proves the "only if" part. For the "if" part we again use Theorem 3 and prove that (4.7), (4.8) imply (4.2), (4.3).

We consider the $s\times s$ matrix M with elements $m_{ij} = b_i\widehat{a}_{ij} + \widehat{b}_j a_{ji} - b_i\widehat{b}_j$ and the vector d whose components are $d_i = b_i - \widehat{b}_i$. Conditions (4.7) and (4.8) (use the equivalent formulation (4.5)) imply that

$$d^T\Phi = 0, \qquad \Phi^T M\Phi = 0, \tag{4.9}$$

where Φ is the matrix of Lemma 4. For irreducible methods this matrix has full rank so that (4.9) implies $d=0$ and $M=0$. This completes the proof of the theorem. □

Remark. The conditions (4.7) and (4.8) only depend on the coefficients $a(t)$ ($t\in TP$) of the P-series (3.4). A recent paper of Calvo & Sanz-Serna [CS93] shows the possibility of a direct proof of Theorem 5 which does not make use of the fact that the coefficients $a(t)$ originate from a partitioned Runge-Kutta method. They have shown that condition (4.8) for $u,v\in T$ is equivalent to the symplecticity of the B-series (2.9) corresponding to a Hamiltonian system.

Separable Hamiltonians. The case where the Hamiltonian function is separable, i.e.,

$$H(p,q) = T(p) + U(q), \tag{4.10}$$

is very interesting because explicit symplectic methods are possible. For instance, the symplectic Euler discretization (1.6) becomes an explicit method (observe that $\partial H/\partial q$ only depends on q). In this situation the condition (4.3) alone implies the symplecticity of the method (see [AS93]). Therefore, the condition (4.7) is no longer necessary. Further, all those elementary differentials vanish which correspond to trees with neighbouring vertices of the same type. Such trees need not be considered. With these modifications in mind, all results of this article hold also for separable Hamiltonians. However, it should be mentioned that the perturbed Hamiltonian (5.7) is not separable.

Nyström Methods. A Hamiltonian system, where

$$H(p,q) = \frac{1}{2}p^T M^{-1} p + U(q) \tag{4.11}$$

with a symmetric matrix M, is given by

$$p' = -U'(q), \qquad q' = M^{-1}p. \tag{4.12}$$

It is equivalent to the second order differential equation

$$q'' = -M^{-1}U'(q). \tag{4.13}$$

Therefore, it is possible to apply special methods for second order systems, e.g., Nyström methods of the form

$$\begin{aligned} Q_i &= q_0 + c_i h M^{-1} p_0 + h^2 \sum_{j=1}^{s} \overline{a}_{ij} k_j', & k_i' &= -M^{-1}U'(Q_i) \\ q_1 &= q_0 + h M^{-1} p_0 + h^2 \sum_{i=1}^{s} \overline{b}_i k_i', & M^{-1}p_1 &= M^{-1}p_0 + h \sum_{i=1}^{s} b_i k_i'. \end{aligned} \tag{4.14}$$

It is known (see [Su89]) that such methods are symplectic, if

$$\overline{b}_i = b_i(1-c_i), \qquad i = 1,\ldots,s \tag{4.15a}$$

$$b_i(\overline{b}_j - \overline{a}_{ij}) = b_j(\overline{b}_i - \overline{a}_{ji}), \qquad i,j = 1,\ldots,s. \tag{4.15b}$$

If we consider the elementary differentials (3.6) for the special system (4.12) we observe that many of them vanish identically. Due to the fact that (4.11) is a separable Hamiltonian, the trees with neighbouring vertices of the same type need not be considered. Furthermore, since $g(p,q) = M^{-1}p$ is a linear function, higher derivatives of g vanish so that the trees to be considered have no ramifications at fat vertices (observe that in [HNW93] the roles of the fat and meagre vertices are interchanged). The so obtained subsets of TP_p and TP_q are denoted by TN_p and TN_q, respectively.

An interesting fact is that the numerical solution (p_1, q_1) of (4.14) can also be written as a P-series (3.4). The coefficients are given by $a(\tau_q) = 1$ and

$$\begin{aligned} a(u) &= \gamma(u) \sum_{i=1}^{s} b_i \Phi_i(u) & &\text{for } u \in TN_p, \\ a(v) &= \gamma(v) \sum_{i=1}^{s} \overline{b}_i \Phi_i(v) & &\text{for } v \in TN_q,\ \rho(v) \geq 2, \end{aligned} \tag{4.16}$$

where $\Phi_i(\tau_p) = 1$,

$$\Phi_i(u) = c_i^r \sum_{j_1,\dots,j_m} \overline{a}_{ij_1}\Phi_{j_1}(v_1)\cdot\ldots\cdot\overline{a}_{ij_m}\Phi_{j_m}(v_m)$$

for $u = {}_p[\overbrace{\tau_q,\dots,\tau_q}^{r}, v_1,\dots,v_m] \in TN_p$ with $\rho(v_i) > 1$ $(i = 1,\dots,m)$, and $\Phi_i(v) = \Phi_i(u_1)$ for $v = {}_q[u_1] \in TN_q$.

We next write the symplecticity conditions (4.15) in terms of the coefficients $a(u)$ and $a(v)$ of (4.16). Multiplying (4.15a) by $\Phi_i(u)$ with $u \in TN_p$ and summing over i yields

$$\sum_i \overline{b}_i\Phi_i(\tau_q\circ u) = \sum_i b_i\Phi_i(u) - \sum_i b_i\Phi_i(u\circ\tau_q).$$

But this is nothing else than formula (4.8) with $v = \tau_q$ because $a(\tau_q) = \gamma(\tau_q) = 1$. If we multipy (4.15b) by $\Phi_i(u)\cdot\Phi_j(\widetilde{u})$ $(u,\widetilde{u}\in TN_p)$ and sum over i and j we obtain

$$\Big(\sum_i b_i\Phi_i(u)\Big)\Big(\sum_j \overline{b}_j\Phi_j({}_q[\widetilde{u}])\Big) - \sum_i b_i\Phi_i(u\circ_q[\widetilde{u}])$$
$$= \Big(\sum_j b_j\Phi_j(\widetilde{u})\Big)\Big(\sum_i \overline{b}_i\Phi_i({}_q[u])\Big) - \sum_i b_j\Phi_j(\widetilde{u}\circ_q[u])$$

so that by (4.16)

$$\frac{a(u)}{\gamma(u)}\cdot\frac{a({}_q[\widetilde{u}])}{\gamma({}_q[\widetilde{u}])} - \frac{a(u\circ_q[\widetilde{u}])}{\gamma(u\circ_q[\widetilde{u}])} = \frac{a(\widetilde{u})}{\gamma(\widetilde{u})}\cdot\frac{a({}_q[u])}{\gamma({}_q[u])} - \frac{a(\widetilde{u}\circ_q[u])}{\gamma(\widetilde{u}\circ_q[u])}. \tag{4.17}$$

But how can we relate this formula to (4.8)? There is a little obstacle because for $u \in TN_p$, $v \in TN_q$ $(\rho(v)\ge 2)$ the product $v\circ u$ is not a tree in TN_q. The idea is to define $a(v\circ u)$ by

$$\frac{a(v\circ u)}{\gamma(v\circ u)} := \frac{a(v)}{\gamma(v)}\cdot\frac{a(u)}{\gamma(u)} - \frac{a(u\circ v)}{\gamma(u\circ v)}. \tag{4.18}$$

But is this coefficient well-defined? For $v = {}_q[\widetilde{u}]$ there are two possibilities to write the tree $v\circ u$ as a product, namely ${}_q[\widetilde{u}]\circ u$ and ${}_q[u]\circ\widetilde{u}$. Due to formula (4.17) both representations lead to the same value so that $a(v\circ u)$ is well-defined by (4.18). Therefore, formula (4.8) holds for all $u\in TN_p$ and $v\in TN_q$ and the results of this article are also valid for Nyström methods.

It should be mentioned that for irreducible Nyström methods both conditions of (4.15) together are equivalent to (4.8). As for separable Hamiltonians (4.10) it is not necessary to consider (4.7). Again, the perturbed Hamiltonian of Theorem 9 below will not be of the special form (4.11).

5. The Perturbed Hamiltonian Function

Theorem 5 above characterizes the symplecticity of a partitioned Runge-Kutta method in terms of the coefficients $a(t)$ of its P-series. Our next result gives conditions for $b(t)$ such that (3.9) represents a Hamiltonian system.

Theorem 6. *The differential equation (3.9) with elementary differentials (formula (3.6)) corresponding to*

$$f(p,q) = -\frac{\partial H}{\partial q}(p,q), \qquad g(p,q) = \frac{\partial H}{\partial p}(p,q) \tag{5.1}$$

is a Hamiltonian system if and only if

$$b(t) \quad \text{is independent of the type of the root of } t, \tag{5.2}$$

$$\frac{b(u\circ v)}{\gamma(u\circ v)} + \frac{b(v\circ u)}{\gamma(v\circ u)} = 0 \qquad \text{for } \; u\in TP_p,\, v\in TP_q. \tag{5.3}$$

We are mainly interested in the "if"-part of this result. For its proof we shall explicitely construct the perturbed Hamiltonian function corresponding to the system (3.9) and begin with some preparations.

5.1 An Equivalence Relation on *TP*. The conditions (4.7), (4.8) and (5.2), (5.3) define certain relations between the coefficients $a(t)$ and $b(t)$, respectively. Trees, which are related by these conditions, will be considered as equivalent.

Definition. We denote by $\sim$ the smallest *equivalence relation* on TP which satisfies the two properties

$u\sim v$ if u and v are identical with the exception of the type of the root;

$u\circ v \sim v\circ u$ for $u\in TP_p$ and $v\in TP_q$.

Equivalent trees of orders up to 3 are grouped together in Fig. 4. All trees of order 2 are equivalent. There are three equivalence classes of trees of order 3. In the case of separable Hamiltonians, equivalence classes for this relation have been considered by [AS93] and are named "bicolor (unrooted) trees".

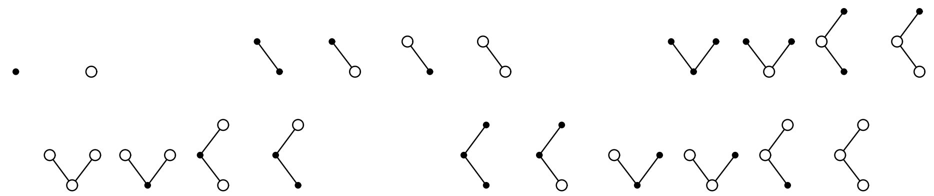


Fig. 4. Equivalent trees

5.2 Elementary Hamiltonians. The perturbed Hamiltonian (see (1.7)) will be composed of socalled "elementary Hamiltonians" (see [SA91], last section) in the same way as the right-hand side of (3.9) is composed of elementary differentials. In the following definition the elementary differentials $F(t)(p,q)$ correspond to (5.1).

Definition. For a given function $H:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ we define the *elementary Hamiltonian* $H(t):\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ by

$$H(\tau_p)(p,q) = H(\tau_q)(p,q) = H(p,q)$$

$$H(t)(p,q) = \frac{\partial^{m+l}H(p,q)}{\partial^m p\,\partial^l q}\Big((-1)^{\delta(u_1)}F(u_1)(p,q),\ldots,(-1)^{\delta(v_1)}F(v_1)(p,q),\ldots\Big)$$

where $t = {}_p[u_1, \ldots, u_m, v_1, \ldots, v_l]$ or $t = {}_q[u_1, \ldots, u_m, v_1, \ldots, v_l]$ with $u_i \in TP_p$ and $v_i \in TP_q$. The integers $\delta(u_i)$ and $\delta(v_i)$ count the number of meagre vertices of u_i and v_i, respectively.

The factors $(-1)^{\delta(u_i)}$ and $(-1)^{\delta(v_i)}$ in the definition of $H(t)$ compensate for the minus sign which is present in the function $f(p,q)$ of (5.1). Examples of elementary Hamiltonians are

$H(\,.\,) = H$, $H($$) = H_p H_q$,

$H($$) = H_{pp} H_q^2$, $H($$) = H_{pq} H_q H_p$, $H($$) = H_{qq} H_p^2$.

We observe that exactly the same expressions appear in formula (1.7).

Lemma 7. *For equivalent trees the elementary Hamiltonians are identical.*

Proof. By definition, the elementary Hamiltonians do not depend on the type of the root. Therefore, all we have to show is that $H(u \circ v)(p,q) = H(v \circ u)(p,q)$ for $u \in TP_p$ and $v \in TP_q$. Let $u = {}_p[u_1, \ldots, u_m]$, $v = {}_q[v_1, \ldots, v_l]$ and denote by μ and λ the numbers of trees in TP_p among $u_1, \ldots, u_m$ and $v_1, \ldots, v_l$, respectively. Omitting the obvious argument (p,q) we then have

$$H(u \circ v) = \pm \sum_{J=1}^{n} \frac{\partial^{m+1} H}{\partial^\mu p \, \partial^{m-\mu} q \, \partial q^J} \Big(F(u_1), \ldots, F(u_m) \Big) \cdot F^J(v) \tag{5.4}$$

where the superscript J denotes the Jth component of a vector. Inserting the expression

$$F^J(v) = \frac{\partial^{l+1} H}{\partial^\lambda p \, \partial^{l-\lambda} q \, \partial p^J} \Big(F(v_1), \ldots, F(v_l) \Big)$$

into (5.4) yields a formula which is completely symmetric with respect to u and v. Since the sign of $H(t)$ is always positive, we must have $H(u \circ v) = H(v \circ u)$. □

For the following results it is convenient to denote the equivalence class of trees by bold letters, i.e., for a tree $t \in TP$ we denote

$$\mathbf{t} = \{ u \in TP \mid u \sim t \}. \tag{5.5}$$

Lemma 8. *For a fixed $t \in TP$ it holds*

$$-\frac{\partial H(t)(p,q)}{\partial q} = \sum_{u \in \mathbf{t} \cap TP_p} \alpha(u) \, \beta(u) \, F(u)(p,q)$$

$$\frac{\partial H(t)(p,q)}{\partial p} = \sum_{v \in \mathbf{t} \cap TP_q} \alpha(v) \, \beta(v) \, F(v)(p,q)$$

where the coefficients $\beta : TP \to \mathbb{R}$ are defined by the relations (5.2), (5.3) (with b replaced by β) and by

$$\rho(t) = \sum_{u \in \mathbf{t} \cap TP_p} (-1)^{\delta(u)+1} \alpha(u) \beta(u) = \sum_{v \in \mathbf{t} \cap TP_q} (-1)^{\delta(v)} \alpha(v) \beta(v). \tag{5.6}$$

The statement is trivial for $t = \tau_p$. For $t =$ we obtain by Leibniz' rule

$$-\frac{\partial H(\;)}{\partial q} = -\frac{\partial}{\partial q}\Big(H_p H_q\Big) = -H_{qp}H_q - H_{qq}H_p = -F(\;) + F(\;)$$
$$\frac{\partial H(\;)}{\partial p} = \frac{\partial}{\partial p}\Big(H_p H_q\Big) = H_{pp}H_q + H_{pq}H_p = -F(\;) + F(\;)$$

so that the statement of Lemma 8 holds with $\beta(\;) = \beta(\;) = -1$ and $\beta(\;) = \beta(\;) = 1$. The proof for the general situation is postponed to Section 6.

We remark that the relations (5.2) and (5.3) determine the coefficients $\beta(u)$ for $u \in \mathbf{t}$ up to a constant factor which is obtained by the normalization (5.6). The coefficients $\beta(u)$ are all different from zero. Their values together with those of $\alpha(t)$ and $\gamma(t)$ are presented in Table 3 for trees with $\rho(t) \leq 3$.

Table 3. Coefficients $\alpha(t)$, $\gamma(t)$, $\beta(t)$

t	$\bullet$ $\circ$									
$\alpha(t)$	1	1	1	1	1	1	1	1	2	1
$\gamma(t)$	1	2	2	3	6	3	6	6	3	6
$\beta(t)$	1	-1	1	1	-2	1	-2	1	$-1/2$	1

5.3 Proof of Theorem 6. a) Under the assumption of (5.2) and (5.3) we shall show that (3.9) is a Hamiltonian system. We search for a perturbed Hamiltonian of the form

$$\widetilde{H}(p,q) = H_1(p,q) + \frac{h}{2!}H_2(p,q) + \ldots + \frac{h^{N-1}}{N!}H_N(p,q), \tag{5.7}$$

where

$$H_k(p,q) = \sum_{t\in R,\rho(t)=k} c(t)\, H(t)(p,q), \tag{5.8}$$

$c(t)$ are real coefficients and R denotes a fixed subset of TP which contains exactly one representative of each equivalence class of $TP/_\sim$. In order that (3.9) is the Hamiltonian system corresponding to (5.7) one needs

$$\begin{aligned} -\sum_{t\in R,\rho(t)=k} c(t)\,\frac{\partial H(t)(p,q)}{\partial q} &= \sum_{u\in TP_p,\rho(u)=k} \alpha(u)\, b(u)\, F(u)(p,q) \\ \sum_{t\in R,\rho(t)=k} c(t)\,\frac{\partial H(t)(p,q)}{\partial p} &= \sum_{v\in TP_q,\rho(v)=k} \alpha(v)\, b(v)\, F(v)(p,q). \end{aligned} \tag{5.9}$$

Inserting the formulas of Lemma 8 this is seen to be true if

$$c(t)\beta(u) = b(u), \qquad c(t)\beta(v) = b(v) \tag{5.10}$$

for $u \in \mathbf{t} \cap TP_p$ and $v \in \mathbf{t} \cap TP_q$. Since the coefficients $\beta(u)$ and $\beta(v)$ are different from zero and both, β and b, satisfy the relations (5.2) and (5.3), the coefficients $c(t)$ are well-defined by (5.10). This completes the proof of the "if"-part of Theorem 6.

b) The "only if"-part is a consequence of the material to be presented below and can be seen as follows: if the system (3.9) is Hamiltonian then its exact solution, which is formally represented by the P-series (3.4), is a symplectic transformation.

Lemma 11 (and its accompanying remark) below then imply that the coefficients $a(t)$ satisfy (4.7) and (4.8). Finally, the statement follows from Lemma 10. $\square$

5.4 The Main Result. We are now in the position to adapt Theorem 2 to Hamiltonian systems and symplectic integration methods.

Theorem 9. *Let $a: TP \to \mathbb{R}$ satisfy the conditions (4.7), (4.8) and assume that the right-hand sides of the Hamiltonian system (1.4) are N times continuously differentiable. The numerical solution p_1, q_1, given by (3.4) with f and g satisfying (5.1), then satisfies*

$$p_1 = \widetilde{p}(x_0 + h) + \mathcal{O}(h^{N+1}), \qquad q_1 = \widetilde{q}(x_0 + h) + \mathcal{O}(h^{N+1}),$$

where $\widetilde{p}(x), \widetilde{q}(x)$ is the exact solution of the perturbed Hamiltonian system

$$\widetilde{p}\,' = -\frac{\partial \widetilde{H}}{\partial q}(\widetilde{p}, \widetilde{q}), \qquad \widetilde{q}\,' = \frac{\partial \widetilde{H}}{\partial p}(\widetilde{p}, \widetilde{q}). \tag{5.11}$$

The new Hamiltonian is given by (5.7) and (5.8). The coefficients $c(t)$ are given by (5.10) with $b(t)$ and $\beta(t)$ defined in Theorem 2 and Lemma 8, respectively.

Proof. It follows from Lemma 10 below that the coefficients $b(t)$ of the perturbed system (3.9) satisfy (5.2) and (5.3). Consequently, this is a Hamiltonian system (Theorem 6). $\square$

Remark. For the special situation where the symplectic Euler method (1.6) is applied to a separable Hamiltonian system, an explicit formula for the perturbed Hamiltonian is given in [Yo93]. The derivation uses the Baker-Campbell-Hausdorff formula and is completely different from our approach.

Lemma 10. *Consider mappings $a: TP \to \mathbb{R}$ and $b: TP \to \mathbb{R}$ which are related by formula (2.12). Then we have*

(4.7) and (4.8) are equivalent to (5.2) and (5.3).

Proof. The equivalence of (4.7) and (5.2) is trivial. The complete proof has originally been obtained by purely combinatorical arguments. The following more elegant proof emerged from discussions with Sebastian Reich on his ideas on backward analysis (see [R93]).

We assume that (4.7) and (4.8) are satisfied and we show by induction on k that

$$\frac{b(u \circ v)}{\gamma(u \circ v)} + \frac{b(v \circ u)}{\gamma(v \circ u)} = 0 \qquad \text{for} \quad \rho(u) + \rho(v) \le k. \tag{5.12}$$

This is clearly satisfied for $k = 0$ (empty statement). Suppose now that (5.12) is satisfied with k replaced by $k-1$ and consider the differential equation

$$\begin{aligned} \widehat{p}\,' &= \sum_{\rho(u)<k} \frac{h^{\rho(u)-1}}{\rho(u)!}\, \alpha(u)\, b(u)\, F(u)(\widehat{p}, \widehat{q}), \\ \widehat{q}\,' &= \sum_{\rho(v)<k} \frac{h^{\rho(v)-1}}{\rho(v)!}\, \alpha(v)\, b(v)\, F(v)(\widehat{p}, \widehat{q}). \end{aligned} \tag{5.13}$$

By the induction hypothesis this is a Hamiltonian system (Theorem 6). Hence

$$
\begin{aligned}
\widehat{p}(x_0+h) &= \sum_{u\in TP_p} \frac{h^{\rho(u)}}{\rho(u)!}\,\alpha(u)\,a_k(u)\,F(u)(p_0,q_0)\\
\widehat{q}(x_0+h) &= \sum_{v\in TP_p} \frac{h^{\rho(v)}}{\rho(v)!}\,\alpha(v)\,a_k(v)\,F(v)(p_0,q_0)
\end{aligned}
\tag{5.14}
$$

is a symplectic transformation and, by Lemma 11 below, the coefficients $a_k(t)$ also satisfy (4.8). The main observation is now that

$$
a_k(t) = \begin{cases} a(t) & \text{for } \rho(t) < k \\ a(t) - b(t) & \text{for } \rho(t) = k. \end{cases} \tag{5.15}
$$

Indeed, the differential equation (5.13) corresponds to (5.14) in exactly the same way as (2.11) to (2.9). Therefore the coefficients $a_k(t)$ are obtained from (2.12) with $b(s)$ $(\rho(s) \geq k)$ replaced by zero and (5.15) follows immediately from (2.13). As a consequence of (5.15) we have $b(u\circ v) = a(u\circ v) - a_k(u\circ v)$ for $\rho(u)+\rho(v)=k$ and the statement (5.12) follows by subtracting equation (4.8) from the analoguous formula for a_k.

The proof of the other direction follows from Theorem 6 and Lemma 11. □

6. Technical Details

In this section we present the missing proofs of the preceeding sections. We begin by introducing the *symmetry* coefficient $\sigma(t)$ (see [Bu87], Sections 140 and 144). It is defined by $\sigma(\tau_p) = \sigma(\tau_q) = 1$ and

$$
\sigma(t) = \sigma(t_1)\cdot\ldots\cdot\sigma(t_m)\cdot\mu_1!\mu_2!\ldots \tag{6.1}
$$

for $t =_p [t_1,\ldots,t_m]$ or $t =_q [t_1,\ldots,t_m]$ with $t_i \in TP$. As in (2.5) the integers $\mu_1,\mu_2,\ldots$ count the numbers of equal trees among $t_1,\ldots,t_m$. From the recursion formulas (2.5), (2.6) and (6.1) one obtains by induction on $\rho(t)$ the relation

$$
\alpha(t)\cdot\gamma(t)\cdot\sigma(t) = \rho(t)! \tag{6.2}
$$

which will be fundamental in the following proofs.

6.1 Characterization of Symplectic P-Series. A complete characterization of symplectic B-series has been given in [CS93]. The following Lemma 11, which is used in the proof of Lemma 10, is one direction of this result (generalized to P-series). In order to keep this article as self-contained as possible we indicate its proof. Its main ideas are the same as those presented in [CS93].

Lemma 11. *Let $F(t)(p,q)$ be the elementary differentials corresponding to (5.1) and suppose that the P-series (3.4) defines a symplectic transformation for general $H(p,q)$. Then the coefficients $a(t)$ satisfy the relation (4.8).*

Proof. a) We fix two trees $u \in TP_p$ and $v \in TP_q$. The idea is to consider a particular (polynomial) Hamiltonian which is formed as follows: attach to the branches of $u\circ v$ the numbers $3,\ldots,r+1$ $(r = \rho(u)+\rho(v))$ such that the branch between the roots of u and v is labelled by 3. Then the Hamiltonian is a sum of as many terms as vertices in the tree. The summand corresponding to a vertex is a product containing the factor p^J (or q^J) if an upward leaving branch "J" is connected with a meagre (or fat) vertex, and the factor q^J (or p^J) if the vertex itself is meagre (or fat) and the branch "J" is leaving downwards. Furthermore, the factors q^2 and p^1 are added to the terms corresponding to the roots of u and v, repectively. For the example of Fig. 5 we have

$$H(p,q) = q^2 q^3 q^4 p^5 + p^1 p^3 p^7 + p^4 p^6 + q^5 + q^6 + q^7. \tag{6.3}$$

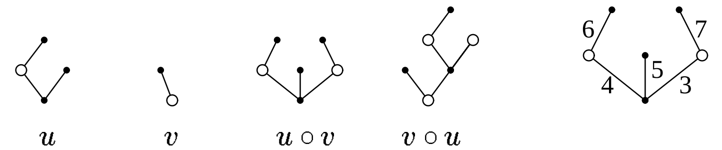


Fig. 5. Construction of the Hamiltonian (6.3)

b) For the Hamiltonian constructed in part (a) we shall prove that the transformation defined by the P-series (3.4) verifies

$$\sum_{J=1}^{r+1} dp_1^J \wedge dq_1^J = \sum_{J=1}^{r+1} dp_0^J \wedge dq_0^J \tag{6.4}$$
$$\pm\, h^{\rho(u)+\rho(v)} \cdot \Big(\frac{a(u\circ v)}{\gamma(u\circ v)} + \frac{a(v\circ u)}{\gamma(v\circ u)} - \frac{a(u)}{\gamma(u)} \cdot \frac{a(v)}{\gamma(v)} \Big) \cdot dp_0^1 \wedge dq_0^2 + \ldots$$

if all expressions are evaluated at $p_0^J = q_0^J = 0$. This, of course, implies the statement of Lemma 11.

Due to the zero initial values only linear terms of $F(t)(p,q)$ in (3.4) give rise to non-vanishing expressions in (6.4). Further we are interested only in the coefficient of $dp_0^1 \wedge dq_0^2$, so that it is sufficient to consider those terms which depend on p^1 and/or q^2. The particular structure of the Hamiltonian of (a) makes the computation of the elementary differentials easy. We observe that the Hamiltonian is independent of q^1. Therefore the first component of $F(t)(p,q)$ vanishes identically for $t \in TP_p$. The second component of $F(t)(p,q)$ ($t \in TP_q$) always starts with a derivative of H with respect to q^2. Therefore it has a non-vanishing term depending on p^1 only if it contains the factor $H_{q^2 q^3 q^4 p^5}$ (for the example of Fig. 5). This in turn implies the presence of factors $H_{p^3\ldots}$, $H_{p^4\ldots}$ and $H_{q^5\ldots}$. Continuing this procedure we find that $F^2(t)(p,q)$ is a linear function of p^1 only for $t = u \circ v$. With similar arguments we see that only the following elementary differentials need to be considered in the transformation (3.4):

$$\begin{aligned}
p_1^2 &= p_0^2 + \frac{h^{\rho(u\circ v)}}{\rho(u\circ v)!}\, \alpha(u\circ v)\, a(u\circ v)\, F^2(u\circ v)(p_0,q_0) + \ldots \\
p_1^3 &= p_0^3 + \frac{h^{\rho(u)}}{\rho(u)!}\, \alpha(u)\, a(u)\, F^3(u)(p_0,q_0) + \ldots \\
q_1^1 &= q_0^1 + \frac{h^{\rho(v\circ u)}}{\rho(v\circ u)!}\, \alpha(v\circ u)\, a(v\circ u)\, F^1(v\circ u)(p_0,q_0) + \ldots \\
q_1^3 &= q_0^3 + \frac{h^{\rho(v)}}{\rho(v)!}\, \alpha(v)\, a(v)\, F^3(v)(p_0,q_0) + \ldots
\end{aligned} \tag{6.5}$$

The values of $F^J(t)(p_0, q_0)$ are computed as

$$
\begin{aligned}
F^2(u \circ v)(p_0, q_0) &= (-1)^{\delta(u \circ v)} \sigma(u \circ v) \cdot p_0^1, \qquad & F^3(u)(p_0, q_0) &= (-1)^{\delta(u)} \sigma(u) \cdot q_0^2, \\
F^1(v \circ u)(p_0, q_0) &= (-1)^{\delta(v \circ u)} \sigma(v \circ u) \cdot q_0^2, \qquad & F^3(v)(p_0, q_0) &= (-1)^{\delta(v)} \sigma(v) \cdot p_0^1.
\end{aligned}
$$

Here the symmetry coefficients σ appear in exactly the same way as in the multi-dimensional Taylor formula. We next insert these expressions for the elementary differentials into (6.5) and use the relation (6.2). The statement (6.4) then follows from the bilinearity and the anti-symmetry of the exterior product. □

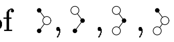
Remarks. a) The hypotheses of Lemma 11 imply the condition (4.7), too. Instead of (6.3) one has to consider a (polynomial) Hamiltonian where the factors q^2 and p^1 are both added to the term corresponding to the root of the considered tree.

b) The reverse of the statement of Lemma 11 (i.e., that (4.7) and (4.8) are also sufficient for the symplecticity of the P-series (3.4)) can be obtained from the results of this article as follows: if the coefficients $a(t)$ satisfy (4.7) and (4.8) then, by Lemma 10, the corresponding coefficients $b(t)$ satisfy (5.2) and (5.3). Hence the differential equation (3.9) is a Hamiltonian system (Theorem 6), so that the transformation (3.4) is necessarily symplectic. This yields an alternative proof for the characterization of symplectic P-series.

c) For P-trees, whose neighbouring vertices are of different type (meagre, fat), the Hamiltonian constructed in the above proof is separable.

6.2 Proof of Lemma 8. We start with an important observation: for two trees $u \in TP_p$ and $v \in TP_q$ the transition from $u \circ v$ to $v \circ u$ (or vice versa) does not change the graph representing these trees, it only changes the position of the root. Therefore, if we fix some vertex of a tree $t \in TP$, it is possible to bring this vertex down to the root just by changing the type of the root and by passing from $u \circ v$ to $v \circ u$ (or from $v \circ u$ to $u \circ v$) without leaving the equivalence class $\mathbf{t}$. Different vertices of t may lead to the same tree u. We thus denote by $\nu(u)$ the number of vertices of t such that $u \in \mathbf{t}$ is obtained by this procedure (this number is independent of the representative t of $\mathbf{t}$). Since the tree t consists of $\rho(t)$ vertices, it holds

$$
\rho(t) = \sum_{u \in \mathbf{t} \cap TP_p} \nu(u) = \sum_{v \in \mathbf{t} \cap TP_q} \nu(v). \tag{6.6}
$$

For all trees of order $\rho(t) \leq 3$ we have $\nu(t) = 1$ with the exception of  for which $\nu(t) = 2$. The use of these coefficients will allow us to prove the statements of Lemma 8.

We fix a tree $t \in TP$ and observe that the elementary Hamiltonian $H(t)(p, q)$ consists of $\rho(t)$ factors corresponding to the $\rho(t)$ vertices of t. We differentiate this expression with the help of Leibniz' rule, but before differentiating the ith factor we bring the corresponding vertex down to the root. Since the elementary Hamiltonian is not changed by this procedure (Lemma 7), the differentiation yields (apart from the sign) an elementary differential $F(u)(p,q)$. We therefore get

$$-\frac{\partial H(t)(p,q)}{\partial q} = \sum_{u\in \mathbf{t}\cap TP_p} \nu(u)\,(-1)^{\delta(u)+1}\,F(u)(p,q)$$
$$\frac{\partial H(t)(p,q)}{\partial p} = \sum_{v\in \mathbf{t}\cap TP_q} \nu(v)\,(-1)^{\delta(v)}\,F(v)(p,q)$$

so that the statement of Lemma 8 holds with

$$\begin{aligned} \beta(u) &= (-1)^{\delta(u)+1}\nu(u)/\alpha(u) \qquad \text{for } u\in \mathbf{t}\cap TP_p \\ \beta(v) &= (-1)^{\delta(v)}\nu(v)/\alpha(v) \qquad \text{for } v\in \mathbf{t}\cap TP_q. \end{aligned} \tag{6.7}$$

Formula (5.6) is an immediate consequence of (6.6). Furthermore, we observe that the number of meagre vertices changes by 1 if the type of the root is changed. Since $\nu(u)$ and $\alpha(u)$ are independent of the type of the root, the same is true for $\beta(u)$. Hence, condition (5.2) is verified and it remains to prove the relation (5.3).

We consider two trees $u = {}_p[u_1,\ldots,u_m]$ and $v = {}_q[v_1,\ldots,v_l]$, for which condition (5.3) becomes

$$\frac{\nu(u\circ v)}{\alpha(u\circ v)\cdot\gamma(u\circ v)} = \frac{\nu(v\circ u)}{\alpha(v\circ u)\cdot\gamma(v\circ u)}$$

and is equivalent to (use formula (6.2))

$$\nu(u\circ v)\cdot\sigma(u\circ v) = \nu(v\circ u)\cdot\sigma(v\circ u). \tag{6.8}$$

This formula can be seen as follows: if there are trees among $v_1,\ldots,v_l$ which equal u, then an exchange of their root with the root of $v\circ u$ yields the tree $u\circ v$. Consequently, the coefficient $\nu(u\circ v)$ is proportional to the number which counts how often u appears among $v_1,\ldots,v_l,u$. But this number can be expressed in terms of the coefficient (6.1) as $\sigma(v\circ u)/\big(\sigma(u)\cdot\sigma(v)\big)$. This proves (6.8) and completes the proof of Lemma 8. □

Acknowledgement. The author thanks G. Wanner for many useful discussions and for his suggestion to use the theory of P-series for extending the results of [Yo93]. He also thanks John Butcher and his collaborators for providing the opportunity to present the results of this paper at the SCADE93 conference in Auckland, and J.M. Sanz-Serna and Ch. Lubich for useful comments.

References

[AS93] L. Abia & J.M. Sanz-Serna: *Partitioned Runge-Kutta methods for separable Hamiltonian problems.* Math. Comput. **60** (1993), p.617-634.

[AF91] S.P. Auerbach & A. Friedman: *Long-time behaviour of numerically computed orbits: small and intermediate timestep analysis of one-dimensional systems.* J. Comp. Phys. **93** (1991), p.189-223.

[Ar89] V.I. Arnol'd: *Mathematical methods of classical mechanics.* Second edition. Springer-Verlag, New York, 1989.

[Bu87] J.C. Butcher: *The numerical analysis of ordinary differential equations. Runge-Kutta and general linear methods.* John Wiley & Sons, Chichester, 1987.

[CS93] M.P. Calvo & J.M. Sanz-Serna: *Canonical B-series.* To appear in Numer. Math.

[GS86] D.F. Griffiths & J.M. Sanz-Serna: *On the scope of the method of modified equations.* SIAM J. Sci. Stat. Comput. **7** (1986), p.994-1008.

[Ha81] E. Hairer: *Order conditions for numerical methods for partitioned ordinary differential equations.* Numer. Math. **36** (1981), p.431-445.

[HNW93] E. Hairer, S.P. Nørsett & G. Wanner: *Solving Ordinary Differential Equations I. Nonstiff Problems.* Second Edition. Springer-Verlag, Berlin, 1993.

[La88] F.M. Lasagni: *Canonical Runge-Kutta methods.* ZAMP **39** (1988), p.952-953.

[MK92] R.S. MacKay: *Some aspects of the dynamics and numerics of Hamiltonian systems.* In: The Dynamics of Numerics and the Numerics of Dynamics, ed. by D.S. Broomhead & A. Iserles, Clarendon Press, Oxford, 1992.

[R93] S. Reich: *Numerical integration of the generalized Euler equations.* To be published.

[SA91] J.M. Sanz-Serna & L. Abia: *Order conditions for canonical Runge-Kutta schemes.* SIAM J. Numer. Anal. **28** (1991), p.1081-1096.

[SS88] J.M. Sanz-Serna: *Runge-Kutta schemes for Hamiltonian systems.* BIT **28** (1988), p.877-883.

[SS92] J.M. Sanz-Serna: *Symplectic integrators for Hamiltonian problems: an overview.* Acta Numerica **1** (1992), p.243-286.

[Su89] Y.B. Suris: *The canonicity of mappings generated by Runge-Kutta type methods when integrating the systems* $\ddot{x} = -\partial U/\partial x$. Zh. Vychisl. Mat. i Mat. Fiz. **29** (1989), p.202-211 (in Russian); same as U.S.S.R. Comput. Maths. Phys. **29**, p.138-144.

[Su90] Y.B. Suris: *Hamiltonian Runge-Kutta type methods and their variational formulation.* Mathematical Simulation **2** (1990), p.78-87 (Russian).

[WH74] R.F. Warming & B.J. Hyett: *The modified equation approach to the stability and accuracy analysis of finite-difference methods.* J. Comp. Phys. **14** (1974), p.159-179.

[Yo93] H. Yoshida: *Recent progress in the theory and application of symplectic integrators.* Celest. Mech. **56** (1993), p.27-43.