MATHEMATICS OF COMPUTATION
VOLUME 60, NUMBER 202
APRIL 1993, PAGES 617-634

PARTITIONED RUNGE-KUTTA METHODS FOR SEPARABLE HAMILTONIAN PROBLEMS

L. ABIA AND J. M. SANZ-SERNA

ABSTRACT. Separable Hamiltonian systems of differential equations have the form $d\mathbf{p}/dt = -\partial H/\partial \mathbf{q}$, $d\mathbf{q}/dt = \partial H/\partial \mathbf{p}$, with a Hamiltonian function H that satisfies $H = T(\mathbf{p}) + V(\mathbf{q})$ (T and V are respectively the kinetic and potential energies). We study the integration of these systems by means of partitioned Runge-Kutta methods, i.e., by means of methods where different Runge-Kutta tableaux are used for the $\mathbf{p}$ and $\mathbf{q}$ equations. We derive a sufficient and "almost" necessary condition for a partitioned Runge-Kutta method to be canonical, i.e., to conserve the symplectic structure of phase space, thereby reproducing the qualitative properties of the Hamiltonian dynamics. We show that the requirement of canonicity operates as a simplifying assumption for the study of the order conditions of the method.

1. INTRODUCTION

The recent literature, both in mathematics and in physics, has shown a great interest in the integration of Hamiltonian systems of differential equations

$$dp^I/dt = -\partial H/\partial q^I, \quad dq^I/dt = \partial H/\partial p^I, \qquad 1 \le I \le N, \tag{1.1}$$

by means of so-called symplectic or canonical methods. An extensive list of references may be found in the survey [10]. In (1.1), $H = H(p^1, \dots, p^N, q^1, \dots, q^N)$ is the Hamiltonian function, a sufficiently smooth real function of $2N$ real variables. The integer N is called the number of degrees of freedom in H. By definition, canonical methods preserve the so-called symplectic structure in $(\mathbf{p}, \mathbf{q})$-space. The flow of (1.1) also preserves the symplectic structure, and this fact actually implies all the properties of the Hamiltonian dynamics [1, 7, 10]. Therefore, canonical methods can be expected to be better suited to the integration of (1.1) than their noncanonical counterparts.

An s-stage Runge-Kutta (RK) method

$$\begin{bmatrix} \mathbf{a} \\ \mathbf{b}^T \end{bmatrix}, \tag{1.2}$$

$\mathbf{a} = [a_{ij}]$, $\mathbf{b} = [b_i]$, is canonical if [6, 9, 13]

$$b_i a_{ij} + b_j a_{ji} - b_i b_j = 0, \qquad 1 \le i, j \le s. \tag{1.3}$$

Received by the editor January 9, 1991 and, in revised form, June 3, 1992.
1991 *Mathematics Subject Classification*. Primary 65L05; Secondary 70H15, 05C05.
Key words and phrases. Hamiltonian systems, symplectic structures, canonical transformations, Runge-Kutta methods, order conditions, bicolor trees, bicolor rooted trees, generating functions.

©1993 American Mathematical Society
0025-5718/93 $1.00 + $.25 per page

Furthermore, for methods without redundant stages, (1.3) is also necessary for canonicity [6]. The well-known RK schemes based on Gauss-Legendre quadrature are canonical [9]. Diagonally implicit methods satisfying (1.3) are derived in [12] and tested in [4]. On the other hand, it is easy to see that (1.3) implies that (1.2) cannot be explicit.

A family of one-step methods that includes *explicit* canonical schemes is given by the so-called partitioned Runge-Kutta (PRK) methods, where different Butcher tableaux are used to advance the p and the q components of (1.1). More precisely, with the notations

$$f^I = -\partial H/\partial q^I\,, \quad g^I = \partial H/\partial p^I\,, \qquad 1 \le I \le N\,, \tag{1.4}$$

a step $(\mathbf{p}, \mathbf{q}) \mapsto (\mathbf{p}^*, \mathbf{q}^*)$ of an s-stage PRK scheme is given by the equations

$$\mathbf{Y}_i = \mathbf{p} + h\sum_{j=1}^{s} a_{ij}\mathbf{k}_j\,, \quad \mathbf{Z}_i = \mathbf{q} + h\sum_{j=1}^{s} A_{ij}\mathbf{l}_j\,, \qquad 1 \le i \le s\,, \tag{1.5}$$

$$\mathbf{p}^* = \mathbf{p} + h\sum_{i=1}^{s} b_i\mathbf{k}_i\,, \qquad \mathbf{q}^* = \mathbf{q} + h\sum_{i=1}^{s} B_i\mathbf{l}_i\,, \tag{1.6}$$

where $\mathbf{k}_i = \mathbf{f}(\mathbf{Y}_i, \mathbf{Z}_i)$, $\mathbf{l}_i = \mathbf{g}(\mathbf{Y}_i, \mathbf{Z}_i)$, $1 \le i \le s$.

In a symbolic form, the method (1.5)–(1.6) can be specified by the tableaux

$$\begin{bmatrix} \mathbf{a} \\ \mathbf{b}^T \end{bmatrix} \begin{bmatrix} \mathbf{A} \\ \mathbf{B}^T \end{bmatrix}. \tag{1.7}$$

In one of the very first papers on canonical integration, Ruth [8] constructed particular examples of methods of the form (1.5)–(1.6) that are canonical when applied to *separable* Hamiltonian systems, i.e., to systems where the Hamiltonian function is of the form

$$H(\mathbf{p}, \mathbf{q}) = T(\mathbf{p}) + V(\mathbf{q})\,. \tag{1.8}$$

(In mechanical applications, T and V correspond to the kinetic and potential energy, respectively.) Ruth did not use the RK formalism and worked with generating functions.

The present paper is devoted to a general study of PRK methods for the numerical integration of Hamiltonian problems. In §2, we give a sufficient condition for a PRK method to be canonical when applied to separable Hamiltonian systems. In §§3 and 4, we show that this sufficient condition operates as a *simplifying assumption* [2, p. 214; 5, p. 203]; i.e., the number of equations for a method to have order p is smaller for methods satisfying this condition than for general methods. In §5 we prove that the sufficient condition for canonicity is "almost" necessary. A numerical example is presented in §6 and the final section contains the conclusions.

Before closing the introduction we would like to mention a convention we have already used: *lower case subscripts* are associated with the stages of the method and run from 1 to s; *upper case superscripts* are associated with the components of the differential system and run from 1 to N.

2. Canonical PRK methods

In this section we investigate under which condition the PRK method (1.5)–(1.6) is canonical. By definition, a method is said to be canonical for a class of

Hamiltonian problems if, for all Hamiltonian functions H in this class, and at each point $(\mathbf{p}_0, \mathbf{q}_0)$ in the domain of definition of H,

$$d\mathbf{p}^* \wedge d\mathbf{q}^* - d\mathbf{p} \wedge d\mathbf{q} = 0, \tag{2.1}$$

whenever the steplength h is small enough for the step $(\mathbf{p}, \mathbf{q}) \mapsto (\mathbf{p}^*, \mathbf{q}^*)$ to be a well-defined mapping in a neighborhood of $(\mathbf{p}_0, \mathbf{q}_0)$.

Exactly the same manipulations used in the proof of Theorem 1 of [9] show that the left-hand side of (2.1) equals

$$h\sum_{i=1}^{s}[b_i d\mathbf{k}_i \wedge d\mathbf{Z}_i + B_i d\mathbf{Y}_i \wedge d\mathbf{l}_i] - h^2 \sum_{i,j=1}^{s} (b_i A_{ij} + B_j a_{ji} - b_i B_j) d\mathbf{k}_i \wedge d\mathbf{l}_j . \tag{2.2}$$

Now it is clear that the second sum in (2.2) will vanish if we demand that

$$b_i A_{ij} + B_j a_{ji} - b_i B_j = 0, \qquad 1 \le i, j \le s, \tag{2.3}$$

a condition which reduces to (1.3) in the case $\mathbf{A} = \mathbf{a}$ and $\mathbf{B} = \mathbf{b}$ (i.e., if the PRK method is actually an RK method). Now, to ensure that the first sum in (2.2) also vanishes, there are two possibilities:

(i) It may be noted that, as proved in [9, p. 880],

$$d\mathbf{k}_i \wedge d\mathbf{Z}_i + d\mathbf{Y}_i \wedge d\mathbf{l}_i = 0, \qquad 1 \le i \le s,$$

so that the condition $b_i = B_i$, $1 \le i \le s$, and (2.3) ensure canonicity.

(ii) Attention may be restricted to *separable* Hamiltonian problems (1.8). Then $\mathbf{f}$ is only a function of $\mathbf{q}$ and

$$d\mathbf{k}_i \wedge d\mathbf{Z}_i = d\mathbf{f}(\mathbf{Z}_i) \wedge d\mathbf{Z}_i = 0, \qquad 1 \le i \le s,$$

because the Jacobian matrix of $\mathbf{f} = -\operatorname{grad} V$ is symmetric and the wedge product is skew-symmetric. Similarly

$$d\mathbf{Y}_i \wedge d\mathbf{l}_i = 0, \qquad 1 \le i \le s,$$

and canonicity follows.

In this paper we restrict our attention to the second of the possibilities above. Let us then state the following theorem.

Theorem 2.1. *Assume that the PRK method* (1.7) *satisfies the condition* (2.3). *Then it is canonical when applied to separable Hamiltonian problems* (1.1), (1.8).

The previous result was announced by one of the present authors at the 1989 London Numerical ODE meeting (see [11]) and has been derived independently by Suris [14]. Actually (2.3) is "almost" necessary for canonicity as we shall later show.

In later developments we shall need the *generating function* of a canonical PRK method [1, 10, 12]. By definition, this is a function $S = S(\mathbf{p}, \mathbf{q}^*; h)$ such that the relations

$$\mathbf{p}^* = \partial S/\partial \mathbf{q}^*, \qquad \mathbf{q} = \partial S/\partial \mathbf{p} \tag{2.4}$$

reproduce the equations (1.6) that define the time-step $(\mathbf{p}, \mathbf{q}) \mapsto (\mathbf{p}^*, \mathbf{q}^*)$ of the method. We have

Theorem 2.2. *Assume that the PRK* (1.7) *satisfies* (2.3) *and is applied to the separable Hamiltonian problem* (1.1), (1.8). *Then the corresponding generating function is given by*

$$S(\mathbf{p}, \mathbf{q}^*; h) = \mathbf{p}^T\mathbf{q}^* - h\sum_i b_i V(\mathbf{Z}_i) - h\sum_i B_i T(\mathbf{Y}_i) + \frac{h^2}{2}\sum_{ij}(B_i a_{ij} - b_j A_{ji} + b_j B_i)\mathbf{f}(\mathbf{Z}_j)^T\mathbf{g}(\mathbf{Y}_i). \tag{2.5}$$

Furthermore,

$$\Gamma(\mathbf{p}, \mathbf{q}^*; h) := \frac{\partial S}{\partial h} = -\sum_i b_i V(\mathbf{Z}_i) - \sum_i B_i T(\mathbf{Y}_i). \tag{2.6}$$

Note that in (2.5)–(2.6) the stages $\mathbf{Z}_j$ and $\mathbf{Y}_i$ are to be interpreted as functions of $\mathbf{p}$, $\mathbf{q}^*$ defined in (1.5)–(1.6). Clearly, the implicit function theorem guarantees that this interpretation is (locally) possible if H is smooth and h small. To prove the theorem, it is sufficient to differentiate (2.5) and to perform some manipulations to arrive at (2.4) and (2.6). The details will not be given.

3. Order conditions

In this section we consider the case where the PRK method (1.5)–(1.6) is applied to a partitioned system of the form

$$d\mathbf{p}/dt = \mathbf{f}(\mathbf{q}), \qquad d\mathbf{q}/dt = \mathbf{g}(\mathbf{p}), \tag{3.1}$$

that *is not assumed to be Hamiltonian*; i.e., we do not suppose that the vector fields $\mathbf{f}$ and $\mathbf{g}$ are gradients of scalar fields $-V$ and T, respectively. For simplicity we still assume that the vectors $\mathbf{p}$ and $\mathbf{q}$ have the same dimension N. (There is no difficulty in extending the material in this section to the case where the dimensions of $\mathbf{p}$ and $\mathbf{q}$ do not coincide. Note also that one can always add dummy equations $dq^I/dt = 0$ or $dp^I/dt = 0$ to ensure that $\mathbf{p}$ and $\mathbf{q}$ have equal dimensions.)

It is well known (see, e.g., [5, §II.14]) how the order p of (1.5)–(1.6) can be characterized in terms of the Butcher arrays in (1.7) via *bicolor rooted trees.* The first column in Figure 1 gives the relevant graphs up to order four. The roots have been marked with crosses. Note that, throughout the paper, we just consider bicolor rooted trees where white (respectively black) vertices have only black (respectively white) sons. This is because, in (3.1), $\mathbf{f}$ only depends on $\mathbf{q}$, while $\mathbf{g}$ only depends on $\mathbf{p}$. For (1.7) to possess order of consistency p, it is necessary and sufficient that for all bicolor rooted trees $\beta\rho\tau$ with order (i.e., number of vertices) $\leq p$

$$\Phi(\beta\rho\tau) = \frac{1}{\gamma(\beta\rho\tau)}, \tag{3.2}$$

where $\Phi(\beta\rho\tau)$ and $\gamma(\beta\rho\tau)$ are, respectively, the *elementary weight* and the *density* of $\beta\rho\tau$.

We recall that the density is an integer-valued function recursively defined as follows: $\gamma(\beta\rho\tau)$ equals the order of $\beta\rho\tau$ times the product of the densities of the bicolor rooted trees that arise when the root of $\beta\rho\tau$ is chopped off, the densities of the bicolor rooted trees with one vertex $\beta\rho\tau_{1,1}$, $\beta\rho\tau_{1,2}$ being

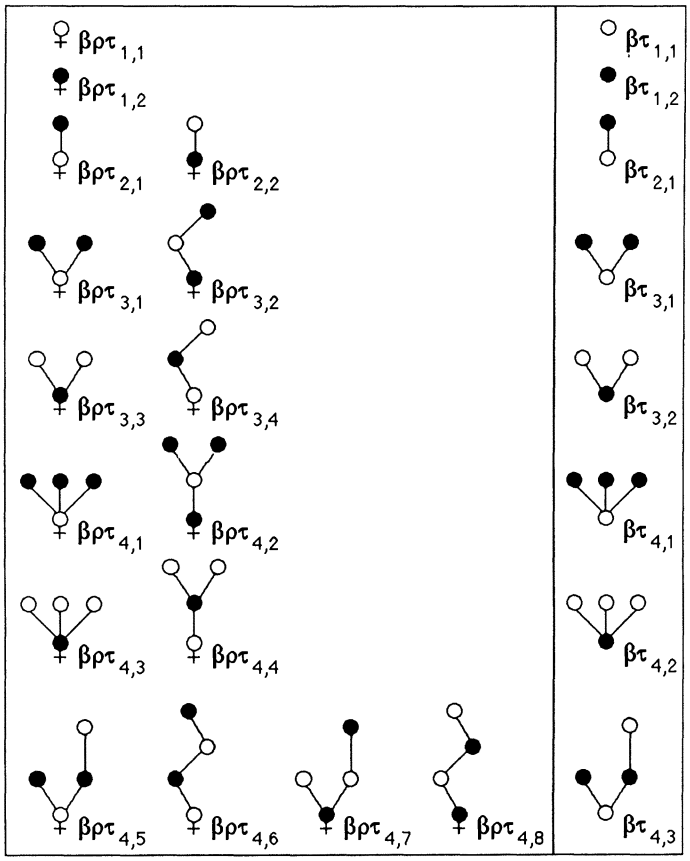

FIGURE 1. Rooted bicoler n-trees and bicolor n-trees, $n = 1, 2, 3, 4$

one. Note that, when computing the density, the coloring of the vertices plays no role; for instance $\beta\rho\tau_{4,6}$ and $\beta\rho\tau_{4,8}$ in Figure 1 have the same density 24. Later, we use the function $\alpha(\beta\rho\tau)$ (number of monotonic labellings [2, §145]). Again, the coloring of the nodes plays no role in the value of this function.

On the other hand, it is well known how to construct the elementary weight corresponding to a given bicolor rooted tree. For instance, for the tree $\beta\rho\tau_{4,5}$ in Figure 1 we have

$$\sum_{ijkl} b_i A_{ij} A_{ik} a_{kl}.$$

In general a white vertex in the graph gives rise to a lower case letter b or a and a black vertex gives rise to an upper case letter B or A. The letters b and B are associated with the root, and the letters a and A with vertices other than the root. There is a summation index associated with each vertex. The index corresponding to the root appears as a subscript of a b or B letter. If a vertex v with index, say, j is son of the vertex with index, say, i, then the letter associated with v has subindices ij.

Let us also write the (partitioned) *elementary differential* associated with the $\beta\rho\tau_{4,5}$. At a point $(\mathbf{p}, \mathbf{q})$ in phase space, this is the N-dimensional vector

whose Ith component is given by

$$\sum_{IJKL} \frac{\partial^2 f^I}{\partial q^J \partial q^K} g^J \frac{\partial g^K}{\partial p^L} f^L$$

(all functions evaluated at $\mathbf{p}$ or $\mathbf{q}$, as appropriate). Now, white (resp. black) vertices are associated with components of $\mathbf{f}$ (resp. $\mathbf{g}$).

The equations (3.2) corresponding to different bicolor rooted trees are not independent if we assume that the PRK method satisfies the canonicity condition (2.3). In other words, (2.3) behave as simplifying assumptions, whereby, to impose order of consistency p, it is not necessary to take into account all bicolor rooted trees with p or fewer vertices. In order to investigate this issue, we need to consider *bicolor* (*unrooted*) *trees*. The bicolor trees of order ≤ 4 are given in the right column of Figure 1. We can define a bicolor tree as an equivalence class consisting of bicolor rooted trees differing only in the choice of the root. In Figure 1 bicolor rooted trees belonging to the same bicolor tree appear in the same row. (Our treatment of the graph theory we need will be sketchy and based on pictorial representations. A rigorous treatment, like that given in [12], would result in too long a paper.)

It will be useful later to observe that, in Figure 1, some bicolor trees, like $\beta\tau_{3,1} - \beta\tau_{3,2}$, appear in pairs, each member of the pair being the photographic negative of the other. However, some bicolor trees are their own photographic negative, like $\beta\tau_{2,1}$ or $\beta\tau_{4,3}$. It is not difficult to see that the latter occurrence takes place if and only if $\beta\tau$ arises from coloring the vertices of what in [12] was called a *superfluous* (uncolored, unrooted) tree.

The main result of this section is then the following.

Theorem 3.1. *Assume that the PRK method* (1.7) *satisfies* (2.3) *and has, at least, oder of consistency* $p-1$, $p \geq 2$, *when applied to all partitioned systems of the form* (3.1). *Let* $\beta\tau$ *be a bicolor tree with* p *vertices. Then the following statements are equivalent*:

(i) *The condition* (3.2) *holds for all bicolor rooted trees* $\beta\rho\tau$ *belonging to* $\beta\tau$.

(ii) *The condition* (3.2) *holds for a bicolor rooted tree* $\beta\rho\tau$ *belonging to* $\beta\tau$.

(iii) *The following homogeneous order condition holds*:

$$\begin{aligned} &\sum_{\substack{\beta\rho\tau\in\beta\tau \\ \beta\rho\tau \text{ has white root}}} \alpha(\beta\rho\tau)\gamma(\beta\rho\tau)\Phi(\beta\rho\tau) \\ &= \sum_{\substack{\beta\rho\tau\in\beta\tau \\ \beta\rho\tau \text{ has black root}}} \alpha(\beta\rho\tau)\gamma(\beta\rho\tau)\Phi(\beta\rho\tau). \end{aligned} \tag{3.3}$$

Proof. Clearly (i) implies (ii). Let us show that (ii) implies (i). Choose two adjacent vertices v and w in $\beta\tau$ and consider four bicolor rooted trees as follows (see Figure 2). We denote by $\beta\rho\tau_v$ (resp. $\beta\rho\tau_w$) the bicolor rooted tree obtained by highlighting in $\beta\tau$ the vertex v (resp. w). We denote by $\beta\rho\tau_V$ and $\beta\rho\tau_W$ the bicolor rooted trees, with roots at v and w, respectively,

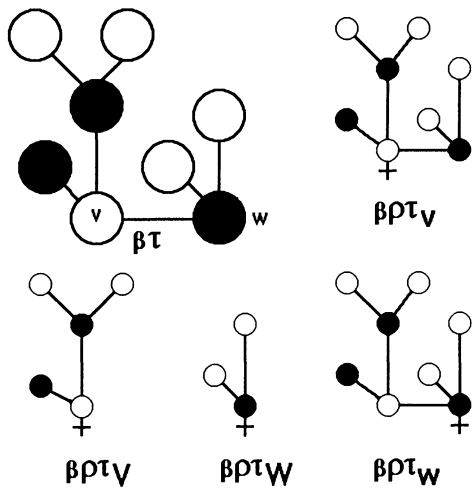

FIGURE 2. The construction for Theorem 3.1

that arise when the edge $[v, w]$ is deleted from $\beta\tau$. On using the argument in the proof of [12, Theorem 5.1], we obtain that

$$\frac{1}{\gamma(\beta\rho\tau_v)} + \frac{1}{\gamma(\beta\rho\tau_w)} = \frac{1}{\gamma(\beta\rho\tau_V)}\frac{1}{\gamma(\beta\rho\tau_W)}$$

and that (2.3) implies

$$\Phi(\beta\rho\tau_v) + \Phi(\beta\rho\tau_w) = \Phi(\beta\rho\tau_V)\Phi(\beta\rho\tau_W).$$

Thus, since the method has order at least $p-1$,

$$\Phi(\beta\rho\tau_v) + \Phi(\beta\rho\tau_w) = \frac{1}{\gamma(\beta\rho\tau_v)} + \frac{1}{\gamma(\beta\rho\tau_w)}, \tag{3.4}$$

so that the equations (3.2) associated with $\beta\rho\tau_v$ and $\beta\rho\tau_w$ are equivalent. This implies that (ii) is sufficient for (i).

The equivalence between (i) and (iii) is established next. We need the identity

$$\sum_{\substack{\beta\rho\tau\in\beta\tau \\ \beta\rho\tau \text{ has white root}}} \alpha(\beta\rho\tau) = \sum_{\substack{\beta\rho\tau\in\beta\tau \\ \beta\rho\tau \text{ has black root}}} \alpha(\beta\rho\tau). \tag{3.5}$$

If the (uncolored) tree τ obtained from $\beta\tau$ by disregarding the coloring of the vertices is superfluous, then (3.5) has been proved in [12, Theorem 4.5]. If τ is superfluous, then (3.5) is trivial. (For instance, for $\beta\tau_{4,3}$ in Figure 1, the sum in the left in (3.5) is $\alpha(\beta\rho\tau_{4,5}) + \alpha(\beta\rho\tau_{4,6})$, the sum in the right is $\alpha(\beta\rho\tau_{4,7}) + \alpha(\beta\rho\tau_{4,8})$. These are the same since the α-function is colorblind.)

Now it is obvious that (3.5) leads to the conclusion that (i) implies (iii). Conversely, assume that $\Phi(\beta\rho\tau)\gamma(\beta\rho\tau) \neq 1$ for some $\beta\rho\tau$ in $\beta\tau$. Let us say that $\Phi\gamma > 1$ and that the root of $\beta\rho\tau$ is black. Then (3.4) reveals that, for bicolor rooted trees belonging to $\beta\rho$, one has $\Phi\gamma > 1$ whenever the root is black and $\Phi\gamma < 1$ whenever the root is white. This and (3.5) show that (3.3) cannot be true. □

Theorem 3.1 enables us to consider, for methods satisfying (2.3), one order condition per bicolor tree, rather than one order condition per rooted bicolor tree. Let us illustrate this point. In general, if a PRK method that is known to have order ≥ 2 is to have order ≥ 3, we must impose the four conditions (cf. Figure 1)

$$\sum_{ijk} b_i A_{ij} A_{ik} = \frac{1}{3}, \qquad \sum_{ijk} B_i a_{ij} A_{jk} = \frac{1}{6}, \tag{3.6}$$

and

$$\sum_{ijk} B_i a_{ij} a_{ik} = \frac{1}{3}, \qquad \sum_{ijk} b_i A_{ij} a_{jk} = \frac{1}{6}. \tag{3.7}$$

If the method satisfies (2.3), the conditions in (3.6) are equivalent, so that one of them can be dispensed with, and, of course, the same applies to the conditions in (3.7). This leaves two conditions. Alternatively, if we wish to use the homogeneous form, for methods satisfying (2.3), both equations in (3.6) are equivalent to

$$3\sum_{ijk} b_i A_{ij} A_{ik} = 6\sum_{ijk} B_i a_{ij} A_{jk},$$

and both equations in (3.7) are equivalent to

$$6\sum_{ijk} b_i A_{ij} a_{jk} = 3\sum_{ijk} B_i a_{ij} a_{ik}.$$

In order to see the number of conditions that remain when (2.3) holds, we proceed as follows. Clearly, the number of bicolor rooted trees of a given order is twice the number of (uncolored) rooted trees of the same order. On the other hand, the number of bicolor trees of a given order p is twice the number of nonsuperfluous p-trees plus the number of superflourous p-trees. On employing the generating functions for the number of rooted trees, trees and superfluous trees (see e.g. [12]), one can now derive the information in Table 1.

TABLE 1

p	Bicolor rooted p-trees	Bicolor p-trees
1	2	2
2	2	1
3	4	2
4	8	3
5	18	6
6	40	10
7	96	22
8	230	42
9	572	94
10	1438	203

Remark 3.1. It is useful to examine what happens in the situation where in (1.7) $\mathbf{a} = \mathbf{A}$ and $\mathbf{b} = \mathbf{B}$; i.e., one is effectively using the RK method (1.2). We begin by noticing that the order of consistency of (1.2) when applied to the restricted class of problems (3.1) is not higher than the classical order, i.e., than the order when applied to general systems of differential equations

$$d\mathbf{y}/dt = \mathbf{f}(\mathbf{y}). \tag{3.8}$$

This can be seen by observing that if $\mathbf{y}(t)$ is a solution of (3.8), then $[\mathbf{y}^T(t), \mathbf{y}^T(t)]^T$ is a solution of the system

$$d\mathbf{p}/dt = \mathbf{f}(\mathbf{q}), \qquad d\mathbf{q}/dt = \mathbf{f}(\mathbf{p}),$$

which is of the form (3.1). Alternatively, it may be noted that, by considering the RK method (1.2) as a particular case of a PRK method, and writing the order conditions as in (3.2), we recover *twice over* the classical order conditions for standard RK methods. (Since each rooted tree can be colored in two different ways, we recover each order condition twice, e.g., $\beta\rho\tau_{4,1}$ and $\beta\rho\tau_{4,3}$ in Figure 1 would both yield the order condition for the (uncolored) bushy rooted tree with four vertices.) Assume now that (1.3) holds, so that we can employ the homogeneous form (3.3) of the order conditions. If $\beta\tau$ arises from coloring a nonsuperfluous tree, then the order condition (3.3) for the photographic negative of $\beta\tau$ just reproduces the order condition for $\beta\tau$. However, if $\beta\tau$ arises from coloring a superfluous tree, then (3.3) is automatically satisfied, as each term in the sum in the right is also present in the sum in the left. We conclude that for RK methods satisfying (1.3) there is one order condition per nonsuperfluous tree, thus recovering Theorem 3.1 of [12].

4. Canonical theory of the order

In this section we assume that the PRK method (1.7) is applied to separable Hamiltonian problems. It is then possible to study the consistency of the method *as applied to this class of problems* via the canonical theory introduced in [12]. However, we should emphasize that the order of consistency of PRK methods is not increased by restricting attention to separable Hamiltonian problems, rather than considering general problems of the form (3.1). This is a consequence of the following result, which implies that, if the method has order p for separable Hamiltonian problems, then (3.2) must hold for all bicolor rooted trees of order $\leq p$ and hence has order p for all problems (3.1).

Theorem 4.1. *For each rooted bicolor tree $\beta\rho\tau_0$ there is a separable Hamiltonian system so that, if $\mathbf{F}(\beta\rho\tau)(\mathbf{0})$ denotes the elementary differential associated with $\beta\rho\tau$ evaluated at $\mathbf{0}$, the following holds true. The first component of (the vector) $\mathbf{F}(\beta\rho\tau_0)(\mathbf{0})$ is $\neq 0$, while the first component of $\mathbf{F}(\beta\rho\tau)(\mathbf{0})$ is 0 for all other bicolor rooted trees.*

Proof. Even though the proof is completely general, the underlying idea is best presented in an example. Let us consider the graph in Figure 3 (next page), where the vertices have been labelled with the root having label 1. We set

$$H = p^2p^3p^4 + p^5p^6 + p^7 + p^8 - q^1q^2q^5q^8 - q^3 - q^4 - q^6q^7.$$

There are as many terms in H as vertices in the graph. The term $p^2p^3p^4$ was introduced because we have a black vertex 2, with sons 3 and 4; the term p^5p^6

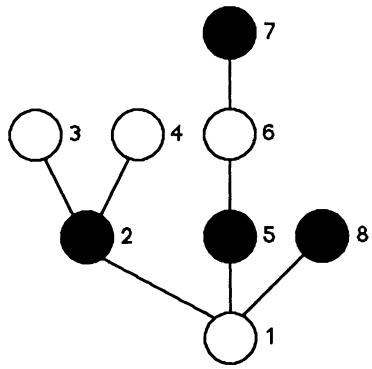

FIGURE 3. Graph for the proof of Theorem 4.1

originates from black vertex 5 with son 6, ..., the term $-q^1q^2q^5q^8$ originates from while vertex 1 with sons 2, 5, and 8, etc. The differential system associated with this Hamiltonian function has $f^1 = q^2q^5q^8$ and $g^1 = 0$. At the origin, all derivatives of these two functions are zero, except for

$$\frac{\partial^3 f^1}{\partial q^2\,\partial q^5\,\partial q^8} = 1\,.$$

Hence, if an elementary differential at $\mathbf{0}$ has nonzero first component, the root of the corresponding graph must be white and have three sons. The iteration of this argument proves that all elementary differentials have 0 first component except that associated with the bicolor rooted tree in Figure 3 (cf. [5, p. 134]). □

Remark 4.1. In Theorem 4.1, the Hamiltonian function can be assumed to be $\mathscr{C}^\infty$, defined in the whole of $\mathscr{R}^{2N}$ and with bounded derivatives of all orders. In fact, it suffices to perturb, away from the origin, the polynomial Hamiltonian constructed in the proof.

The order theory of one-step canonical methods introduced in [12] implies that, if (1.7) is consistent, then it has order $p \geq 2$ if and only if the function Γ defined in (2.6) satisfies, at $h = 0$,

$$\left.\frac{\partial^{n-1}\Gamma}{\partial h^{n-1}}\right|_{\mathbf{p},\mathbf{q}} = 0, \qquad 2 \leq n \leq p\,. \tag{4.1}$$

Here the notation in the partial derivative means that, once Γ has been found as a function of $\mathbf{p}$, $\mathbf{q}^*$, and h, following the prescription (2.6), we should see $\mathbf{q}^*$ as a function of $\mathbf{p}$, $\mathbf{q}$, and h, and differentiate with respect to h, with $\mathbf{p}$, $\mathbf{q}$ fixed. To explicitly compute the partial derivative in (4.1), we use Faa di Bruno's formula [5, p. 148] to express $\partial^{n-1}\Gamma/\partial h^{n-1}|_{\mathbf{p},\mathbf{q}}$ in terms of partial derivatives of T and V and of derivatives of the stages $\mathbf{Y}_i$, $\mathbf{Z}_i$ with respect to h. Then the latter derivatives can be written in terms of the elementary differentials. The result is

$$\left.\frac{\partial^{n-1}\Gamma}{\partial h^{n-1}}\right|_{\mathbf{p},\mathbf{q}} = -\frac{1}{n}\sum_{\beta\rho\tau \text{ has order } n} \alpha(\beta\rho\tau)\gamma(\beta\rho\tau)\Phi(\beta\rho\tau)\theta(\beta\rho\tau)\,, \tag{4.2}$$

where $\theta(\beta\rho\tau)$ is a real-valued function of $\mathbf{p}$ and $\mathbf{q}$. For the bicolor rooted

trees with three vertices (see Figure 1), the explicit expressions are

$$\theta(\beta\rho\tau_{3,1}) = \sum_{IJ} \frac{\partial^2 V}{\partial q^I \partial q^J} g^I g^J, \tag{4.3}$$

$$\theta(\beta\rho\tau_{3,2}) = \sum_{IJ} \frac{\partial T}{\partial p^I} \frac{\partial f^I}{\partial q^J} g^J, \tag{4.4}$$

$$\theta(\beta\rho\tau_{3,3}) = \sum_{IJ} \frac{\partial^2 T}{\partial p^I \partial p^J} f^I f^J, \tag{4.5}$$

$$\theta(\beta\rho\tau_{3,4}) = \sum_{IJ} \frac{\partial V}{\partial q^I} \frac{\partial g^I}{\partial p^J} f^J. \tag{4.6}$$

In general, there are as many summation indices as vertices other than the root. A white (resp. black) root with r sons $I, J, \ldots$ introduces a factor $\partial^r V/\partial q^I \partial q^J \cdots$ (resp. $\partial^r T/\partial p^I \partial p^J \cdots$). A white (resp. black) vertex, different from the root, with label I and r sons labelled $J, K, \ldots$ brings in a factor $\partial^r f^I/\partial q^J \partial q^K \partial \cdots$ (resp. $\partial^r g^I/\partial p^J \partial p^K \partial \cdots$).

Now the key observation is that, for any two bicolor rooted tree $\beta\rho\tau$ and $\beta\rho\tau'$ belonging to the same bicolor tree,

$$\theta(\beta\rho\tau) = \pm\theta(\beta\rho\tau'), \tag{4.7}$$

where, if both graphs have roots of the same (resp. different) color, the $+$ (resp. $-$) sign has to be taken. To prove (4.7), it is sufficient to use, in the definition of θ, the relations $\mathbf{f} = -\operatorname{grad} V$, $\mathbf{g} = \operatorname{grad} T$ in order to eliminate the partial derivatives of V, T. For instance, from (4.3)–(4.4), we may write

$$\theta(\beta\rho\tau_{3,1}) = -\sum_{IJ} \frac{\partial f^I}{\partial q^J} g^I g^J,$$

$$\theta(\beta\rho\tau_{3,2}) = \sum_{IJ} g^I \frac{\partial f^I}{\partial q^J} g^J.$$

Once (4.7) is available, (4.2) may be rewritten as

$$\left.\frac{\partial^{n-1}\Gamma}{\partial h^{n-1}}\right|_{\mathbf{p},\mathbf{q}} = -\frac{1}{n} \sum_{\beta\tau \text{ order } n} \left[\sum_{\substack{\beta\rho\tau\in\beta\tau \\ \text{white root}}} \alpha(\beta\rho\tau)\gamma(\beta\rho\tau)\Phi(\beta\rho\tau) - \sum_{\substack{\beta\rho\tau\in\beta\tau \\ \text{black root}}} \alpha(\beta\rho\tau)\gamma(\beta\rho\tau)\Phi(\beta\rho\tau) \right] \hat{\theta}(\beta\tau), \tag{4.8}$$

where $\hat{\theta}(\beta\tau)$ denotes the common value of θ for the white-rooted bicolor rooted trees in $\beta\tau$. It follows that, for a consistent method that satisfies (2.3) to have order $p > 1$, it is *sufficient* that the homogeneous order condition (3.3) holds for all bicolor trees $\beta\tau$ with $\leq p$ vertices. This, of course, agrees with Theorem 3.1. Furthermore, this condition for order p is also necessary, because the $\hat{\theta}$'s are independent. More precisely, we have the following result.

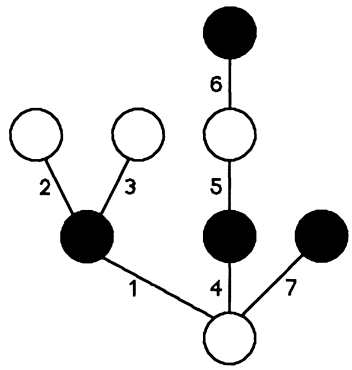

FIGURE 4. Graph for the proof of Theorem 4.2

Theorem 4.2. *With the notations above, to each bicolor tree* $\beta\tau_0$ *there corresponds a separable Hamiltonian such that, at* $\mathbf{p} = \mathbf{q} = \mathbf{0}$, $\hat{\theta}(\beta\tau_0) \neq 0$, *while* $\hat{\theta}(\beta\tau) = 0$ *for all bicolor trees* $\beta\tau \neq \beta\tau_0$.

Proof. It is useful to begin by expressing $\hat{\theta}(\beta\tau)$ in terms of derivatives of V and T (i.e., to use the relations $\mathbf{f} = -\operatorname{grad} V$, $\mathbf{g} = \operatorname{grad} T$ to eliminate the explicit occurrence of the derivatives of the components of $\mathbf{f}$ and $\mathbf{g}$). For instance, from (4.3),

$$\hat{\theta}(\beta\tau_{3,1}) = \theta(\beta\rho\tau_{3,1}) = \sum_{IJ} \frac{\partial^2 V}{\partial q^I \partial q^J} \frac{\partial T}{\partial p^I} \frac{\partial T}{\partial p^J}.$$

In general, $\hat{\theta}(\beta\tau)$ is given by a sum with as many summation indices as *edges* are in the graph. To each edge there corresponds a summation index. A white (resp. black) vertex entered by r edges $I, J, \ldots$ contributes with a factor $\partial^r V/\partial q^I \partial q^J \cdots$ (resp. $\partial^r T/\partial p^I \partial p^J \cdots$). Finally, the sum should be taken with a plus (resp. minus) sign if the number of white vertices is odd (resp. even).

After this preparation we are ready for the proof. Again we consider the graph in Figure 3, but now we ignore the root and label the edges rather than the vertices (see Figure 4). The corresponding Hamiltonian is

$$H = q^1 q^4 q^7 + q^2 + q^3 + q^5 q^6 + p^1 p^2 p^3 + p^4 p^5 + p^6 + p^7.$$

There are as many terms as vertices. A white (resp. black) vertex entered by the edges labelled $I, J, \ldots$ contributes with a term $q^I q^J \cdots$ (resp. $p^I p^J \cdots$). In order to verify that H does in fact meet the requirements, one follows a recursive procedure similar to that used in the proof of Theorem 4.2. □

Remark 4.2. Once more, it is useful to look at the situation where (1.7) reduces to a single RK method (1.2). For trees $\beta\tau$ that are their own photographic negative, the bracket in (4.8) vanishes: each term in the sum for white roots is cancelled by a term in the sum for black roots. Next consider a tree $\beta\tau$ with photographic negative $\beta\tau' \neq \beta\tau$. If $\beta\rho\tau \in \beta\tau$ has white root and photographic negative $\beta\rho\tau'$, then $\alpha(\beta\rho\tau)\gamma(\beta\rho\tau)\Phi(\beta\rho\tau) = \alpha(\beta\rho\tau')\gamma(\beta\rho\tau')\Phi(\beta\rho\tau')$. Thus we can group together the contributions of $\beta\rho\tau$ and $\beta\rho\tau'$ to (4.8). The

combined contribution is easily seen to be

$$\alpha(\beta\rho\tau)\gamma(\beta\rho\tau)\Phi(\beta\rho\tau)[\hat{\theta}(\beta\tau) - \hat{\theta}(\beta\tau')]. \tag{4.9}$$

By carrying out a similar pairing in all terms, we recover, for the particular case of separable Hamiltonians, formula (7.7) of [12]. The expression in brackets in (4.9) is what in [12] was called an elementary canonical differential. The θ notation used in this paper is slightly different from that used in [12].

5. Necessary conditions for canonicity

In this section we study whether (2.3) is necessary for canonicity. In general it is not: Each of the methods

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{5.1}$$

and

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{5.2}$$

satisfies (2.3) and is canonical. One step of length h of the method

$$\begin{bmatrix} 0 & 0 \\ 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}, \qquad \begin{bmatrix} 1/2 & 0 \\ 1/2 & 0 \\ 1/2 & 1/2 \end{bmatrix} \tag{5.3}$$

is just a step of length $h/2$ of (5.1) followed by a step of length $h/2$ of (5.2). Hence, (5.3) is also canonical. Now, for (5.3), $\mathbf{Z}_1 \equiv \mathbf{Z}_2$ and hence (5.3) gives the same results as

$$\begin{bmatrix} 0 & 0 \\ 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}, \qquad \begin{bmatrix} 1/2 & 0 \\ 1/2 & 0 \\ \lambda & 1-\lambda \end{bmatrix}, \tag{5.4}$$

where λ is any real number. Thus (5.4) is canonical. However it does not satisfy (2.3) if $\lambda \neq 1/2$. This shows that, if one deals with methods for which two of the stages are equivalent, then canonicity does not imply (2.3). In the remainder of the section we focus our attention on PRK methods without equivalent stages.

For standard RK methods (1.2), the notion of equivalent stages is carefully studied in [2, §383]. This material can easily been extended to the case of PRK methods; full details will not be given here. It is sufficient to mention that two stages $\mathbf{Y}_i$ and $\mathbf{Y}_j$ (or $\mathbf{Z}_k$ and $\mathbf{Z}_l$) of the method (1.7) are said to be equivalent if, to each smooth problem (3.1) and each initial point $(\mathbf{p}, \mathbf{q})$ there corresponds a value h_0 such that $\mathbf{Y}_i = \mathbf{Y}_j$ (or $\mathbf{Z}_i = \mathbf{Z}_j$), for $h < h_0$. The equivalence of two stages can be characterized either in terms of stage elementary weights or in terms of the Butcher tableaux (1.7) (see [2, §383]). The following result holds.

Lemma 5.1. *Assume that the PRK method* (1.7) *has no pair of equivalent stages. Then, there exists a* $\mathscr{C}^\infty$ *separable Hamiltonian, defined in the whole of* $\mathscr{R}^{2N}$ *and with bounded partial derivatives of all orders, such that, at* $\mathbf{p} = \mathbf{q} = \mathbf{0}$, *for all sufficiently small* $h > 0$ *and all* i, j, $i \neq j$,

$$\mathbf{Y}_i \neq \mathbf{Y}_j \quad \textit{and} \quad \mathbf{Z}_i \neq \mathbf{Z}_j.$$

Proof. By characterizing the equivalence of stages in terms of stage elementary weights, as in [2, Theorem 383B], we see that there is a white-rooted bicolor

rooted tree for which the elementary weights of the first and second stages are different. We use Theorem 4.1 and Remark 4.1 to construct a $\mathscr{C}^\infty$ separable Hamiltonian with bounded derivatives for which, for h small and at the origin in phase space, $\mathbf{Y}_1 \neq \mathbf{Y}_2$ (cf. the proof of Lemma 383A in [2]). In a like manner one can construct Hamiltonians that make $\mathbf{Z}_1 \neq \mathbf{Z}_2$, $\mathbf{Y}_1 \neq \mathbf{Y}_3$, etc. By juxtaposing all the corresponding Hamiltonian systems, we find a Hamiltonian system for which all $\mathbf{Y}$-stages are pairwise different and all $\mathbf{Z}$-stages are also pairwise different. □

In the remainder of the section, we look at the class $\mathscr{H}^\infty$ of all $\mathscr{C}^\infty$ separable Hamiltonian functions H defined in the whole of $\mathscr{R}^{2N}$, with bounded partial derivatives of all orders. The corresponding differential systems satisfy a global Lipschitz condition. Hence, given a PRK method (1.7) and $H \in \mathscr{H}^\infty$, there is a constant σ, depending only on the method and on the number of degrees of freedom N, such that, if Λ denotes the lowest upper bound of the modulus of the second partial derivatives of H, then the equations (1.5) have a unique solution for

$$h\Lambda < \sigma . \tag{5.5}$$

It follows that, in these circumstances, $(\mathbf{p}, \mathbf{q}) \mapsto (\mathbf{p}^*, \mathbf{q}^*)$ is a well-defined smooth transformation $\mathscr{T}_{H,h}$ in $\mathscr{R}^{2N}$. The main result in this section is the following theorem.

Theorem 5.1. *If, for a given PRK method without equivalent stages, $\mathscr{T}_{H,h}$ is a canonical transformation for each $H \in \mathscr{H}^\infty$ and each h satisfying* (5.5), *then* (2.3) *holds.*

Proof. Our technique of proof is similar to that used by Lasagni for the case of standard RK methods (1.2) and general Hamiltonian problems. (It should be mentioned that Lasagni's proof does not appear in the published note [6] and is known to us through a kind personal communication.)

We begin by applying Lemma 5.1 to construct an N-degrees of freedom separable Hamiltonian $H_0 = T_0 + V_0 \in \mathscr{H}^\infty$ such that, for h small and initial value $(\mathbf{0}, \mathbf{0})$, the stages are pairwise different. Next, for $1 \le i \le s$, let $M_{T,i} = [t_i^{IJ}]$ and $M_{V,i} = [v_i^{IJ}]$ be $N \times N$ real symmetric matrices with

$$|t_i^{IJ}| \le 1, \quad |v_i^{IJ}| \le 1, \qquad 1 \le i \le s, \ 1 \le I, J \le N. \tag{5.6}$$

We consider the quadratic Hamiltonian functions

$$\begin{aligned} H_i(\mathbf{p}, \mathbf{q}) &= T_i(\mathbf{p}) + V_i(\mathbf{q}), \\ T_i(\mathbf{p}) &= T_0 + \frac{\partial T_0}{\partial \mathbf{p}}(\mathbf{p} - \mathbf{Y}_i) + \frac{1}{2}(\mathbf{p} - \mathbf{Y}_i)^T M_{T,i}(\mathbf{p} - \mathbf{Y}_i), \\ V_i(\mathbf{q}) &= V_0 + \frac{\partial V_0}{\partial \mathbf{q}}(\mathbf{q} - \mathbf{Z}_i) + \frac{1}{2}(\mathbf{q} - \mathbf{Z}_i)^T M_{V,i}(\mathbf{q} - \mathbf{Z}_i), \end{aligned} \tag{5.7}$$

where T_0, $\partial T_0/\partial \mathbf{p}$ are evaluated at $\mathbf{Y}_i$, and V_0, $\partial V_0/\partial \mathbf{q}$ are evaluated at $\mathbf{Z}_i$. Thus, in the neighborhood of $(\mathbf{Y}_i, \mathbf{Z}_i)$, H_i, $1 \le i \le s$, and H_0 differ in quadratic terms. Note that H_i depends on h through $\mathbf{Y}_i$ and $\mathbf{Z}_i$, but this dependence is not reflected in the notation. Finally, let φ be a $\mathscr{C}^\infty$ real function defined in $\mathscr{R}^N$ which is $\equiv 1$ in a neighborhood of the origin and $\equiv 0$

outside a second, larger neighborhood of the origin, and set, for $0 < \varepsilon < 1$,

$$
\begin{aligned}
H(\mathbf{p}, \mathbf{q}) = &\left(1 - \sum_{i=1}^{s} \varphi(\varepsilon^{-1}(\mathbf{p} - \mathbf{Y}_i))\right) T_0(\mathbf{p}) + \sum_{i=1}^{s} \varphi(\varepsilon^{-1}(\mathbf{p} - \mathbf{Y}_i)) T_i(\mathbf{p}) \\
&+ \left(1 - \sum_{i=1}^{s} \varphi(\varepsilon^{-1}(\mathbf{q} - \mathbf{Z}_i))\right) V_0(\mathbf{q}) + \sum_{i=1}^{s} \varphi(\varepsilon^{-1}(\mathbf{q} - \mathbf{Z}_i)) V_i(\mathbf{q}).
\end{aligned} \tag{5.8}
$$

Clearly, the function in (5.8) is in $\mathscr{H}^\infty$. Let us now determine, for h small, $\varepsilon = \varepsilon(h)$ small enough to ensure that $H \equiv H_i$ in a neighborhood of $(\mathbf{Y}_i, \mathbf{Z}_i)$, $1 \le i \le s$; this is possible because the points $\mathbf{Y}_i$ are pairwise different and the same is true for the points $\mathbf{Z}_i$. Then, let us reduce the value of h for (5.5) to hold. This is possible because, as $\varepsilon \to 0$, the second derivatives of H can be proved to remain bounded. (For instance, when forming the matrix of second derivatives of (5.8), one gets the term $\varepsilon^{-2} D^2 \varphi(\varepsilon^{-1}(\mathbf{p} - \mathbf{Y}_i))[T_i - T_0]$. According to a previous remark, the term in brackets is of the order of ε^2 and this offsets the effect of the factor ε^{-2}.) By assumption, for H in (5.8) and the value of h we have found, (2.1) holds. Therefore, the second sum in (2.2) vanishes (the first sum is 0, because H is separable). In order to evaluate this second sum, it is sufficient to observe that the PRK stages for the systems with Hamiltonians H and H_0 are the same, because of the unicity of the solution of the implicit equations guaranteed by (5.5). We can then write

$$
\begin{aligned}
0 &= \sum_{i,j} (b_i A_{ij} + B_j a_{ji} - b_i B_j) d\mathbf{k}_i \wedge d\mathbf{l}_j \\
&= \sum_{i,j} \sum_{I,J,K} (b_i A_{ij} + B_j a_{ji} - b_i B_j) v_i^{I,J} t_j^{I,K} dq^J \wedge dp^K,
\end{aligned}
$$

and the independence of $dq^J \wedge dp^K$ ensures that, for each J and K,

$$
\sum_{i,j} (b_i A_{ij} + B_j a_{ji} - b_i B_j) v_i^{I,J} t_j^{I,K} = 0.
$$

Since $v_i^{I,J}$ and $t_j^{I,K}$ are arbitrary, subject to (5.6), we have proved that (2.3) must hold. □

6. A NUMERICAL ILLUSTRATION

As mentioned in the introduction, the main potential advantage of canonical PRK methods over canonical RK methods is that the former may be explicit. In fact, Ruth [8] considered methods of the form

$$
\begin{bmatrix} c_1 & 0 & 0 & \cdots & 0 \\ c_1 & c_2 & 0 & \cdots & 0 \\ c_1 & c_2 & c_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_1 & c_2 & c_3 & \cdots & c_s \\ c_1 & c_2 & c_3 & \cdots & c_s \end{bmatrix}, \qquad \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ d_1 & 0 & 0 & \cdots & 0 \\ d_1 & d_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_1 & d_2 & d_3 & \cdots & 0 \\ d_1 & d_2 & d_3 & \cdots & d_s \end{bmatrix}. \tag{6.1}
$$

These satisfy (2.3) and are effectively explicit for partitioned problems of the form (3.1). (The equation $\mathbf{Y}_1 = \mathbf{p} + hc_1\mathbf{k}_1$ has not to be solved for $\mathbf{Y}_1$ because $\mathbf{k}_1$ depends only on $\mathbf{Z}_1$, etc.) In [8], Ruth studies the cases $s = 1, 2, 3$.

When $s = 3$ there are six parameters c_i and d_i and, since (2.3) holds, five equations are required to achieve order 3. This leaves a one-parameter family of three-stage, third-order methods of the form (6.1). The family may be described as follows. We consider in the (d_1, d_2)-plane the cubic curve

$$12d_1d_2(d_1+d_2) - 9(d_1^2 + 3d_1d_2 + d_2^2) + 12(d_1+d_2) - 4 = 0. \tag{6.2}$$

Note that for a given value of d_1, there are zero, one or two (real) points (d_1, d_2) on (6.2), whose d_2 coordinate can be found by solving a quadratic equation. Then, to each (real) point on the curve (6.2), there corresponds a method with

$$\begin{aligned} c_1 &= \frac{3d_1 + 3d_2 - 2}{6d_1d_2}, & c_3 &= \frac{-3d_1 + 2}{6d_2(d_1+d_2)}, \\ c_1 &= 1 - c_2 - c_3, & d_3 &= 1 - d_1 - d_2. \end{aligned} \tag{6.3}$$

In [11], one of the present authors suggested to use in (6.2) the value $d_1 = 0.91966152$ (a root of $12z^4 - 24z^2 + 16z - 3 = 0$). For this choice one gets the symmetry $c_i = d_{4-i}$, $i = 1, 2, 3$.

If we take a step of length $h/2$ with (6.1), followed by a step of length $h/2$ of the adjoint method of (6.1) (see [5, §II.8]), we obtain an explicit method with $2s$ stages, which is also canonical, because the inverse of a canonical transformation is canonical (cf. method (5.2) above). In the case $s = 3$ the resulting tableaux are

(6.4)

$$\begin{bmatrix} c_1/2 & 0 & 0 & 0 & 0 & 0 \\ c_1/2 & c_2/2 & 0 & 0 & 0 & 0 \\ c_1/2 & c_2/2 & c_3/2 & 0 & 0 & 0 \\ c_1/2 & c_2/2 & c_3/2 & 0 & 0 & 0 \\ c_1/2 & c_2/2 & c_3/2 & c_3/2 & 0 & 0 \\ c_1/2 & c_2/2 & c_3/2 & c_3/2 & c_2/2 & 0 \\ c_1/2 & c_2/2 & c_3/2 & c_3/2 & c_2/2 & c_1/2 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ d_1/2 & 0 & 0 & 0 & 0 & 0 \\ d_1/2 & d_2/2 & 0 & 0 & 0 & 0 \\ d_1/2 & d_2/2 & d_3/2 & d_3/2 & 0 & 0 \\ d_1/2 & d_2/2 & d_3/2 & d_3/2 & d_2/2 & 0 \\ d_1/2 & d_2/2 & d_3/2 & d_3/2 & d_2/2 & d_1/2 \\ d_1/2 & d_2/2 & d_3/2 & d_3/2 & d_2/2 & d_1/2 \end{bmatrix}.$$

If we start with a three-stage, third-order method (6.1)–(6.3), then (6.4) has order *four*. This comes about because, on the one hand, (6.4), being a composition of order-3 methods, has at least order three, and, on the other hand, the order of (6.4) is even, as the method is its own adjoint. Even though (6.4) has six stages, it only requires *five* evaluations of $\mathbf{f}$ and $\mathbf{g}$ per step. To see this, it is enough to realize that in (6.4) (i) the stage $\mathbf{Y}_4$ duplicates the stage $\mathbf{Y}_3$, which saves a $\mathbf{g}$ evaluation, and (ii) the $\mathbf{f}$ evaluation at $\mathbf{Z}_6$ in the current stage provides the first evaluation of $\mathbf{f}$ to be used at the next step. We have thus constructed a one-parameter family of fourth-order, five evaluation per step, explicit canonical integrators (6.2)–(6.4). We employ the member of this family corresponding to the value of the free parameter d_1 given above.

We now present a simple numerical example. We emphasize that our purpose here is to illustrate the scope of symplectic integrators, rather than to offer definite conclusions as to the relative merits of symplectic methods when compared with their standard, nonsymplectic counterparts. The reader is referred to [3] for more serious numerical comparisons using variable-step software.

We consider the well-known Kepler problem, given by $N = 2$, $T = ((p^1)^2 + (p^2)^2)/2$, $V = -1/\sqrt{((q^1)^2 + (q^2)^2)}$, with initial condition

$$p^1 = 0, \quad p^2 = \sqrt{\frac{1+e}{1-e}}, \quad q^1 = 1 - e, \quad q^2 = 0.$$

Here e, the eccentricity of the orbit, is a parameter that in the experiments is given the value 0.3. The solution is 2π-periodic. The problem is integrated for 10,000 complete periods and the errors are measured at the final time in the Euclidean norm of $\mathscr{R}^4$. We have employed the symplectic PRK described above with stepsizes $2\pi/128$, $2\pi/256$, $2\pi/512$, $2\pi/1024$ and, as a reference algorithm, the classical fourth-order RK method. The latter was run with smaller stepsizes: $2\pi/160$, $2\pi/320$, $2\pi/640$, $2\pi/1280$; this choice makes up for the fact that the reference algorithm uses four function evaluations per step and the symplectic algorithm requires five function evaluations per step. The results given in Table 2 clearly show the advantages of the symplectic algorithm.

TABLE 2

	Classical RK		Symplectic	
Function Evaluations	h	Error	h	Error
6,400,000	$2\pi/160$	********	$2\pi/128$	$0.43E-02$
12,800,000	$2\pi/320$	$0.54E+00$	$2\pi/256$	$0.27E-03$
25,600,000	$2\pi/640$	$0.17E-01$	$2\pi/512$	$0.17E-04$
51,200,000	$2\pi/1280$	$0.53E-03$	$2\pi/1024$	$0.11E-05$

7. CONCLUSIONS

We have considered PRK methods, i.e., methods where the Runge-Kutta formula is used to integrate some components of the solution and a different Runge-Kutta formula is used to integrate the remaining components. For PRK schemes without redundant stages, we have given a necessary and sufficient condition that the coefficients should satisfy for the method to be symplectic when applied to Hamiltonian problems with a separable Hamiltonian $H = T(\mathbf{p}) + V(\mathbf{q})$. It has been shown how to write the order conditions for symplectic PRK methods. These conditions are a subset of the order conditions for general PRK methods. An explicit, 4th-order, symplectic PRK formula has been constructed. A simple numerical example has clearly shown the potential advantages of symplectic integrators.

ACKNOWLEDGMENT

This research has been supported by "Junta de Castilla y León" under project 1031-89 and by Dirección General de Investigación Científica y Técnica under project PB89-0351.

BIBLIOGRAPHY

1. V. I. Arnold, *Mathematical methods of classical mechanics*, 2nd ed., Springer, Berlin, 1989.
2. J. C. Butcher, *The numerical analysis of ordinary differential equations*, Wiley, Chichester, 1987.
3. M. P. Calvo and J. M. Sanz-Serna, *The development of variable-step symplectic integrators, with application to the two-body problem*, SIAM J. Sci. Statist. Comput. (to appear).

4. J. de Frutos and J. M. Sanz-Serna, *An easily implementable fourth-order method for the time integration of wave problems*, J. Comput. Phys. **103** (1992), 160–168.
5. E. Hairer, S. P. Nørsett, and G. Wanner, *Solving ordinary differential equations*. I, *Nonstiff problems*, Springer, Berlin, 1987.
6. F. M. Lasagni, *Canonical Runge-Kutta methods*, Z. Angew. Math. Phys. **39** (1988), 952–953.
7. R. S. Mackay, *Some aspects of the dynamics and numerics of Hamiltonian systems*, The Dynamics of Numerics and the Numerics of Dynamics (D. S. Broomhead and A. Iserles, eds.), Clarendon Press, Oxford, 1992, pp. 137–193.
8. R. Ruth, *A canonical integration technique*, IEEE Trans. Nuclear Sci. **30** (1984), 2669–2671.
9. J. M. Sanz-Serna, *Runge-Kutta schemes for Hamiltonian systems*, BIT **28** (1988), 877–883.
10. ——, *Symplectic integrators for Hamiltonian problems: an overview*, Acta Numerica **1** (1992), 243–286.
11. ——, *The numerical integration of Hamiltonian systems*, Computational Ordinary Differential Equations (J. R. Cash and I. Gladwell, eds.), Clarendon Press, Oxford, 1992, pp. 437–449.
12. J. M. Sanz-Serna and L. Abia, *Order conditions for canonical Runge-Kutta schemes*, SIAM J. Numer. Anal. **28** (1991), 1081–1096.
13. Y. B. Suris, *Canonical transformations generated by methods of Runge-Kutta type for the numerical integration of the system* $x'' = -\partial U/\partial x$, Zh. Vychisl. Mat. i Mat. Fiz. **29** (1987), 202–211. (Russian)
14. ——, *Hamiltonian methods of Runge-Kutta type and their variational interpretation*, Mat. Model. **2** (1990), 78–87. (Russian).

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

E-mail address: abia@cpd.uva.es

E-mail address: sanzserna@cpd.uva.es