SIAM J. NUMER. ANAL.
Vol. 33, No. 1, pp. 368-387, February 1996

019

SYMPLECTIC PARTITIONED RUNGE–KUTTA METHODS FOR CONSTRAINED HAMILTONIAN SYSTEMS*

LAURENT JAY†

Abstract. This article deals with the numerical treatment of Hamiltonian systems with holonomic constraints. A class of partitioned Runge–Kutta methods, consisting of the couples of s-stage Lobatto IIIA and Lobatto IIIB methods, has been discovered to solve these problems efficiently. These methods are symplectic, preserve all underlying constraints, and are superconvergent with order $2s-2$. For separable Hamiltonians of the form $H(q, p) = \frac{1}{2}p^T M^{-1} p + U(q)$ the Rattle algorithm based on the Verlet method was up to now the only known symplectic method preserving the constraints. In fact this method turns out to be equivalent to the 2-stage Lobatto IIIA–IIIB method of order 2. Numerical examples have been performed which illustrate the theoretical results.

Key words. differential-algebraic equations, Hamiltonian systems, index 3, Lobatto methods, mechanical systems, partitioned Runge–Kutta methods, symplecticity

AMS subject classifications. 65L06

1. Introduction.

Hamiltonian problems arise in a lot of applications where dissipative forces can be neglected, such as mechanical systems, astronomy, electrodynamics, molecular dynamics, plasma physics, fluid dynamics, etc. The *Hamiltonian system* of differential equations associated with the *Hamiltonian* $H(q, p)$ (a real function supposed sufficiently smooth) is given by

$$\dot{q} = H_p^T(q, p)\,, \qquad \dot{p} = -H_q^T(q, p) \tag{1.1}$$

where $q = (q^1, \ldots, q^n)^T \in \mathbf{R}^n$ are the *generalized coordinates* and $p = (p^1, \ldots, p^n)^T \in \mathbf{R}^n$ the *generalized momenta*. The flow generated in the phase space $\mathbf{R}^n \times \mathbf{R}^n$ of (q, p) by these equations (1.1) is known to be *symplectic*, i.e., the differential 2-form

$$\omega^2 = \sum_{k=1}^{n} dq^k \wedge dp^k \qquad \text{is preserved,} \tag{1.2}$$

implying that all differential $2d$-forms

$$\underbrace{\omega^2 \wedge \ldots \wedge \omega^2}_{d \text{ times}} \qquad \text{for} \quad d = 1, \ldots, n \tag{1.3}$$

are also conserved ($d = n$ corresponds to the $2n$-form *volume*). Another specific feature of such systems is that the Hamiltonian along a solution $(q(t), p(t))$ to (1.1) passing through (q_0, p_0) at t_0 remains invariant, i.e.,

$$H(q(t), p(t)) = H(q_0, p_0) \qquad \text{for all } t. \tag{1.4}$$

Hamiltonian systems also possess numerous other specific properties (see [Ar89, Part III] and [MK92]). Unfortunately, most numerical methods applied to (1.1) do not maintain the above two properties (1.2) and (1.4). Various authors ([SS88], [La88], [Sur89], [Y90], [Sun92], [Sun93], [ASS93] among others) have identified or constructed *symplectic schemes*, i.e., methods maintaining (1.2). For an overview on symplectic integrators we refer to [SS92] and [HNW93, §II.16].

*Received by the editors June 15, 1993; accepted for publication (in revised form) May 10, 1994.

†Department of Computer Science, University of Minnesota, 4-192 EE/CS Bldg., 200 Union Street S.E., Minneapolis, MN 55455-0159 (na.ljay@na-net.ornl.gov).

In this article we consider Hamiltonian systems with holonomic constraints. Such problems form a particular class of semi-explicit differential-algebraic equations (DAEs) of index 3. We present a very efficient class of partitioned Runge–Kutta (PRK) methods for the solution of these problems. It consists of the couples of s-stage Lobatto IIIA and Lobatto IIIB methods. These methods combine three attractive properties:

—symplecticity, as seen with the RK methods Gauss and Lobatto IIIS (see [Sun92], [Sun93], and [Cha90]);

—the fact that the numerical solution can be naturally projected onto the manifolds where the exact solution lies (see [AP91], [AP92], [HW91, §VI.7], [J94], and [LS94] for similar ideas), without loss of symplecticity;

—superconvergence, a property shared by stiffly accurate RK methods such as Radau IIA and Lobatto IIIC (see [HLR89, pp. 18–19, 86], [J93b], [HJ93], [J94], and [J95]).

The importance of symplecticity in numerical integration, especially for long-time computations, is nowadays underlined by a sort of "backward analysis" by interpreting the numerical solution as the *exact* solution of a nearby *perturbed* Hamiltonian system (see [SS92]). In [H94] Hairer recently proved the general result that *all* symplectic methods whose numerical solution is a (partitioned) P-series possess this property. An extension of this result to Hamiltonian systems with holonomic constraints and the numerical methods treated in this article is likely. This fact is corroborated by the numerical observations of §5.

A typical example of a constrained Hamiltonian system is given by the pendulum equations. Using the cartesian coordinates $q=(x, z)^T$ for the description of the position of the pendulum, the holonomic constraint on the length ℓ of the rod of the pendulum is

$$0=\sqrt{x^2+z^2}-\ell\,. \tag{1.5}$$

The kinetic energy T and the potential energy U of the system are given, respectively, by

$$T(\dot q)=\frac{m}{2}\left(\dot x^2+\dot z^2\right),\qquad U(q)=mgz \tag{1.6}$$

where g is the gravitational constant. The Lagrangian of the system is $L(q, \dot q)=T(\dot q)-U(q)$ and the generalized momenta are $p=(p_x, p_z)^T=L_{\dot q}^T(q, \dot q)$ leading to

$$p_x=m\dot x,\qquad p_z=m\dot z\,. \tag{1.7}$$

The Hamiltonian $H=T+U$ can be expressed by

$$H(q, p)=\frac{1}{2m}\left(p_x^2+p_z^2\right)+mgz\,, \tag{1.8}$$

and the Hamilton equations of motion become

$$\dot x=\frac{p_x}{m},\qquad \dot z=\frac{p_z}{m},\qquad \dot p_x=-\frac{x}{\ell}\lambda,\qquad \dot p_z=-mg-\frac{z}{\ell}\lambda\,. \tag{1.9}$$

One differentiation of (1.5) implies that

$$0=xp_x+zp_z\,, \tag{1.10}$$

and another one permits us to obtain

$$\lambda=\frac{1}{\ell}\left(\frac{1}{m}\left(p_x^2+p_z^2\right)-mgz\right)\,. \tag{1.11}$$

Thus the differential-algebraic system (1.5) with (1.9) is of differential index 3 (see [BCP89, Chap. 2], [HLR89, p. 1], and [HW91, §VI.5] for various *index* definitions).

This paper is organized in five sections. In §2 we give some basic definitions and results related to symplectic PRK methods. Section 3 deals with the application of PRK methods to Hamiltonian systems with holonomic constraints. Convergence results for a specific class of PRK methods, comprising the symplectic Lobatto IIIA–IIIB schemes, are then stated in §4. Finally, §5 includes some numerical experiments illustrating the theoretical results.

2. Symplectic PRK methods for Hamiltonian systems. Hamiltonian systems (1.1) are intrinsically split into two parts; therefore the use of *PRK methods* is very natural.

DEFINITION 2.1. *One step of an s-stage PRK method applied to* (1.1), *with stepsize h and initial values* (q_0, p_0) *at* t_0 *reads*

$$q_1 = q_0 + h\sum_{i=1}^{s} b_i k_i\,, \qquad p_1 = p_0 + h\sum_{i=1}^{s} \widehat{b}_i \ell_i \tag{2.1a}$$

where

$$k_i = H_p^T(Q_i, P_i)\,, \qquad \ell_i = -H_q^T(Q_i, P_i)\,, \tag{2.1b}$$

and the internal stages are given by

$$Q_i = q_0 + h\sum_{j=1}^{s} a_{ij} k_j\,, \qquad P_i = p_0 + h\sum_{j=1}^{s} \widehat{a}_{ij} \ell_j\,. \tag{2.1c}$$

From now on we use the notation (A, b, c) for the RK coefficients where $A := (a_{ij})_{i,j=1}^{s}$, $b := (b_1, \ldots, b_s)^T$, $c := (c_1, \ldots, c_s)^T$, and similarly for $(\widehat{A}, \widehat{b}, \widehat{c})$ (the coefficients c_i and $\widehat{c}_i$ enter into the definition of PRK methods for nonautonomous problems).

For a PRK method the symplecticity condition (1.2) is expressed by

$$\sum_{k=1}^{n} dq_1^k \wedge dp_1^k = \sum_{k=1}^{n} dq_0^k \wedge dp_0^k\,, \tag{2.2}$$

and symplectic PRK schemes can be characterized as follows:

THEOREM 2.1 [H94]. *If the coefficients of an s-stage PRK method* (2.1) *satisfy*

$$b_i = \widehat{b}_i \qquad \textit{for} \quad i = 1, \ldots, s\,, \tag{2.3a}$$

$$b_i\widehat{a}_{ij} + \widehat{b}_j a_{ji} - b_i\widehat{b}_j = 0 \qquad \textit{for} \quad i = 1, \ldots, s\,, \quad j = 1, \ldots, s\,, \tag{2.3b}$$

then the PRK method is symplectic.

If the PRK method is irreducible, then the conditions (2.3) *are also necessary for symplecticity.*

Remarks.

(1) For *separable Hamiltonians* $H(q, p) = T(p) + U(q)$, the first condition (2.3a) can be omitted (see [ASS93]).

(2) For a proof of the sufficiency of the conditions (2.3) see [SS88], [La88], [Sur89], and [Sun92].

(3) For *irreducible* PRK methods, i.e., methods without equivalent stages (see [Bu87, §383], [H94, §4], or the definition of *S-irreducibility* in [HW91, p. 200]), a way of showing the necessity of (2.3) is to extend the characterization of canonical B-series of [CSS94] to

(partitioned) P-series (see [H94, Lemma 11]) and to apply the proof of [H94, Thm. 5]. For separable Hamiltonians an alternative proof of the necessity of (2.3b) is given in [ASS93].

DEFINITION 2.2. *The local error of a PRK method* (2.1) *is defined by*

$$\delta q_h(t_0) = q_1 - q(t_0+h)\,, \qquad \delta p_h(t_0) = p_1 - p(t_0+h) \tag{2.4}$$

where $(q(t), p(t))$ *is the exact solution of* (1.1) *passing through* (q_0, p_0) *at* t_0.

Considering the elegant W-transformation of Hairer and Wanner (see [HW91, §IV.5]), it is possible to construct high order symplectic PRK methods starting from known RK methods as follows.

THEOREM 2.2 [Sun92]. *Suppose that a RK method with coefficients* a_{ij}, $b_i \neq 0$, *and distinct* c_i, *satisfies the following simplifying assumptions:*

$$B(p): \quad \sum_{i=1}^{s} b_i c_i^{k-1} = \frac{1}{k} \quad \text{for} \quad k = 1, \ldots, p\,,$$

$$C(q): \quad \sum_{j=1}^{s} a_{ij} c_j^{k-1} = \frac{c_i^k}{k} \quad \text{for} \quad i = 1, \ldots, s\,, \quad k = 1, \ldots, q\,,$$

$$D(r): \quad \sum_{i=1}^{s} b_i c_i^{k-1} a_{ij} = \frac{b_j}{k}\left(1 - c_j^k\right) \quad \text{for} \quad j = 1, \ldots, s\,, \quad k = 1, \ldots, r\,;$$

then the PRK method (2.1) *with coefficients* $\widehat{b}_i := b_i$, $\widehat{c}_i := c_i$, *and* $\widehat{a}_{ij} := b_j(1 - a_{ji}/b_i)$ *is symplectic and satisfies*

$$\delta q_h(t_0) = \mathcal{O}(h^{\eta+1})\,, \qquad \delta p_h(t_0) = \mathcal{O}(h^{\eta+1}) \tag{2.5}$$

with an order $\eta = \min(p, 2q+2, 2r+2, q+r+1)$.

Remarks.

(1) With the help of the W-transformation it can be shown that the RK method $(\widehat{A}, b, c)$ satisfies $\widehat{C}(r)$ and $\widehat{D}(q)$ (see [HW91, §IV.5] and [Sun92]) where the notation $\widehat{C}(\widehat{q})$, $\widehat{D}(\widehat{r})$ is self-evident.

(2) The simplifying assumptions $C(1)$ and $D(1)$ (which are equivalent to $\widehat{C}(1)$ by the symplecticity conditions (2.3)) ensure here that $c_i = \sum_{j=1}^{s} a_{ij}$ and $c_i = \sum_{j=1}^{s} \widehat{a}_{ij}$, respectively. This implies some simplifications when deriving the order conditions of PRK methods applied to nonautonomous problems: in this case the order conditions reduce to those of the autonomous case (see also [HNW93, p. 134]). An example of a RK method violating one of these assumptions is given by the 2-stage Lobatto IIIA method, namely the *trapezoidal rule*, which satisfies $B(2)$, $C(2)$, but not $D(1)$. Another example consists in its dual symplectic method, the 2-stage Lobatto IIIB method, which satisfies $B(2)$, $D(2)$, but not $C(1)$.

(3) The symplecticity conditions (2.3), acting as simplifying assumptions, introduce a reduction of the number of order conditions (see [ASS93]).

Examples of symplectic PRK methods are given in [Sun92]. In this paper we focus our attention on PRK methods adapted to the situation where holonomic constraints are encountered. In this context the couples of s-stage Lobatto IIIA methods for (A, b, c) and Lobatto IIIB methods for $(\widehat{A}, b, c)$ turn out to be of main interest. These PRK methods satisfy the simplifying assumptions $B(2s-2)$, $C(s)$, $D(s-2)$, $\widehat{C}(s-2)$, and $\widehat{D}(s)$. Concerning the coefficients of these methods, the weights c_i of Lobatto quadratures are given by $c_1 = 0$, $c_s = 1$, and the remaining c_i for $i = 2, \ldots, s-1$ are the roots of the polynomial of degree $s-2$ $P_{s-2}^{(1,1)}(2x-1)$ where

$$P_{s-2}^{(1,1)}(y) = \text{Const} \cdot \frac{1}{(y^2-1)} \frac{d^{s-2}}{dy^{s-2}} \left((y^2-1)^{s-1}\right) \tag{2.6}$$

TABLE 2.1
Coefficients of the 2-stage Lobatto IIIA–IIIB method of order 2.

0	0	0
1	1/2	1/2
	1/2	1/2

0	1/2	0
1	1/2	0
	1/2	1/2

TABLE 2.2
Coefficients of the 3-stage Lobatto IIIA-IIIB method of order 4.

0	0	0	0
1/2	5/24	1/3	−1/24
1	1/6	2/3	1/6
	1/6	2/3	1/6

0	1/6	−1/6	0
1/2	1/6	1/3	0
1	1/6	5/6	0
	1/6	2/3	1/6

is a Jacobi polynomial. The coefficients $b_j = a_{sj}$ and a_{ij} can be computed, for example, by the use of $C(s)$, and the coefficients $\widehat{a}_{ij}$, as in Theorem 2.2 or with the help of $\widehat{D}(s)$. The Butcher tableaux of the 2- and 3-stage Lobatto IIIA–IIIB methods are given in Tables 2.1 and 2.2, respectively. For separable Hamiltonians the 2-stage method can be applied explicitly. The coefficients of the 4-stage Lobatto IIIA–IIIB method can be found in [HW91, p. 80]. We list also below the weights c_i of the 11-stage Lobatto method of order 20:

$$
\begin{aligned}
&c_1 = 0, \quad c_2 = \frac{1}{2}-\delta_2, \quad c_3 = \frac{1}{2}-\delta_3, \quad c_4 = \frac{1}{2}-\delta_4, \quad c_5 = \frac{1}{2}-\delta_5, \quad c_6 = \frac{1}{2}, \\
&c_7 = \frac{1}{2}+\delta_5, \quad c_8 = \frac{1}{2}+\delta_4, \quad c_9 = \frac{1}{2}+\delta_3, \quad c_{10} = \frac{1}{2}+\delta_2, \quad c_{11} = 1, \\
&\delta_2 = \frac{1}{2}\sqrt{\frac{1}{2}(-u+a_1)+a_3}, \qquad \delta_3 = \frac{1}{2}\sqrt{\frac{1}{2}(-u-a_1)+a_3}, \\
&\delta_4 = \frac{1}{2}\sqrt{\frac{1}{2}(u+a_2)+a_3}, \qquad \delta_5 = \frac{1}{2}\sqrt{\frac{1}{2}(u-a_2)+a_3}, \qquad (2.7) \\
&a_1 = \sqrt{v-4\alpha}, \quad a_2 = \sqrt{v-4\beta}, \quad a_3 = \frac{9}{19}, \quad \alpha = d_1-d_2, \quad \beta = d_1+d_2, \\
&d_1 = \frac{1}{2}(e_1+v), \quad d_2 = \frac{1}{2}\sqrt{(e_1+v)^2-4e_2}, \quad e_1 = -\frac{1080}{6137}, \quad e_2 = \frac{96912}{28800941}, \\
&u = \sqrt{v}, \quad v = \frac{32}{323c}\cdot\cos\left(\frac{1}{3}\arccos\left(\frac{3}{2c}\right)\right)+\frac{720}{6137}, \quad c = \sqrt{\frac{13}{42}}.
\end{aligned}
$$

Due to their symmetry, Lobatto schemes are often used for the solution of boundary value problems (see [A85] and [AMR88]). The analysis of the application of Lobatto IIIA methods to semi-explicit index-2 DAEs in Hessenberg form is given in [J93a].

3. Hamiltonian systems with holonomic constraints and PRK methods. Mechanical systems where dissipative forces can be neglected can be expressed as Hamiltonian systems. Their Hamiltonian is of the form $H = T + U$ where T represents the *kinetic energy* and U the *potential energy*. Usually the equations of motion are not written with a Hamiltonian formalism, but in an Euler–Lagrange formulation (see [HW91, §§VI.5 and VI.9]). Our aim here is to study Hamiltonian systems with $m < n$ *holonomic* constraints $g^i(q) = 0$ $(i = 1, \ldots, m)$ (see [Cho92, Chap. 2] or [Ar89, §17]). A Lagrange-type variational principle exists if the constraints are holonomic (see [Cho92, §4.2.2]). Applying this principle to $\widehat{H}(q, p) = H(q, p) + \lambda^T g(q)$, we arrive at

$$\text{(3.1a,b,c)} \qquad \dot{q} = H_p^T(q, p)\,, \qquad \dot{p} = -H_q^T(q, p) - G^T(q)\lambda\,, \qquad 0 = g(q)$$

where $G(q) := g_q(q)$. The variables λ^i $(i = 1, \ldots, m)$ are the *Lagrange multipliers*. Differentiating (3.1c) twice, as in the example given in the Introduction, we obtain the following additional constraints (omitting the obvious function arguments):

$$\text{(3.1d,e)} \qquad 0 = GH_p^T\,, \qquad 0 = G_q(H_p^T, H_p^T) + GH_{pq}^T H_p^T - GH_{pp}^T H_q^T - GH_{pp}^T G^T\lambda\,.$$

Initial values for the problem (3.1a,b,c) have to be *consistent*, i.e., they must satisfy (3.1c,d,e). From now on we suppose that G is of full row rank m and that we have an *optical Hamiltonian system* (see [MK92, p. 140]), meaning that H_{pp} is a strictly positive definite matrix. From these hypotheses it follows that the matrix $GH_{pp}^T G^T$ is invertible; hence we get from (3.1e)

$$\text{(3.1f)} \qquad \lambda(q, p) = \left(GH_{pp}^T G^T\right)^{-1} \left(G_q(H_p^T, H_p^T) + GH_{pq}^T H_p^T - GH_{pp}^T H_q^T\right)(q, p)\,,$$

and thus the original system (3.1a,b,c) is of differential index 3. This explicit relation (3.1f) for λ introduced in (3.1b) defines the *standard underlying ODE* (3.1a,b) which is not a Hamiltonian system in general. All equations (3.1a,b,c,d,e) form an *overdetermined* system of *differential-algebraic equations* (*ODAEs*) (see [FL89] and [PP92]) of index 1. A standard analysis shows that on the $2(n-m)$-dimensional manifold

$$\text{(3.2)} \qquad V = \left\{(q, p) \in \mathbf{R}^n \times \mathbf{R}^n \mid 0 = g(q)\,,\ 0 = G(q)H_p^T(q, p)\right\}\,,$$

the flow generated by the equations (3.1) is symplectic (see also [HNW93, §I.14]).

Disregarding the property (1.4), the ideal properties for a numerical method would be to be symplectic, to have a numerical solution remaining on the manifold V, and to have a high order of convergence occuring with minimal computational work. The Gauss methods applied to (3.1a,b,c) are symplectic, but they have the disadvantages that the numerical solution does not satisfy the constraints (3.1c,d,e) and that a poor (or even no) convergence occurs (see [HJ93]). Even if projections are effected they are not superconvergent, and the symplecticity property is destroyed. To our knowledge the *Rattle algorithm*, a method of order 2 due to Andersen (see [An83]) and based on the *Verlet method*, is the only known symplectic method preserving the constraints which has been proposed in the literature for separable Hamiltonians of the form $H(q, p) = \frac{1}{2}p^T M^{-1} p + U(q)$ with M a constant positive definite matrix (see [LS94] and [SBO92]). For such Hamiltonians two different approaches have also been derived in [LR94]. The first one is the reduction of (3.1a,b,c) to a Hamiltonian state-space form via a parametrization of the constraints (3.1c). The second one is the construction of an unconstrained Hamiltonian system which preserves the constraint manifold V and whose flow reduces to the flow of (3.1) along this manifold. We also mention the independent and simultaneous investigation by Reich on constrained Hamiltonian systems in [R93] where high order symplectic methods preserving the constraints are constructed by composition of a first

order method (see also [Y90]). For short-time computations, a nonsymplectic alternative is to integrate the standard underlying ODE and to frequently project the numerical solution onto the manifolds (3.1c,d,e) (see also [AP93]).

Here we turn our interest to PRK methods.

DEFINITION 3.1. *The application of an s-stage PRK method to the equations* (3.1a,b,c) *reads*

$$q_1 = q_0 + h\sum_{i=1}^{s} b_i k_i\,, \qquad p_1 = p_0 + h\sum_{i=1}^{s} \widehat{b}_i \ell_i \tag{3.3a}$$

where

$$k_i = H_p^T(Q_i, P_i)\,, \qquad \ell_i = -H_q^T(Q_i, P_i) - G^T(Q_i)\Lambda_i\,, \qquad 0 = g(Q_i)\,, \tag{3.3b}$$

and the internal stages are given by

$$Q_i = q_0 + h\sum_{j=1}^{s} a_{ij} k_j\,, \qquad P_i = p_0 + h\sum_{j=1}^{s} \widehat{a}_{ij} \ell_j\,. \tag{3.3c}$$

Remark. The existence and uniqueness of a solution to these equations is not guaranteed without some assumptions on the coefficients.

Let us start by supposing that we have a locally unique solution to this system (3.3). Our aim now is to check the symplecticity condition (2.2) for PRK methods with coefficients satisfying (2.3). It is no surprise that we have the following result.

THEOREM 3.1. *If the coefficients of the PRK method* (3.3) *satisfy* (2.3), *and if* (q_1, p_1) *are uniquely determined, then the numerical flow* $(q_0, p_0) \mapsto (q_1, p_1)$ *is symplectic.*

Proof. This proof is inspired by the calculations of [SS88], [SS92], and [HNW93, Thm. II.16.6]. We neglect the equations $0 = g(Q_i)$ for the moment. Using (3.3a) and the bilinearity of the wedge product "$\wedge$" we compute for $k = 1, \ldots, n$

$$dq_1^k \wedge dp_1^k - dq_0^k \wedge dp_0^k = h\sum_{i=1}^{s} b_i dk_i^k \wedge dp_0^k + h\sum_{j=1}^{s} \widehat{b}_j dq_0^k \wedge d\ell_j^k + h^2 \sum_{i,j=1}^{s} b_i \widehat{b}_j dk_i^k \wedge d\ell_j^k\,. \tag{3.4}$$

We then replace the differentials dq_0^k, dp_0^k with the help of (3.3c) and we obtain

$$\begin{aligned} &dq_1^k \wedge dp_1^k - dq_0^k \wedge dp_0^k \\ &= h\sum_{i=1}^{s} b_i dk_i^k \wedge dP_i^k + h\sum_{j=1}^{s} \widehat{b}_j dQ_j^k \wedge d\ell_j^k - h^2 \sum_{i,j=1}^{s} (b_i \widehat{a}_{ij} + \widehat{b}_j a_{ji} - b_i \widehat{b}_j) dk_i^k \wedge d\ell_j^k\,. \end{aligned} \tag{3.5}$$

An easy calculation shows that

$$\sum_{k=1}^{n} dk_i^k \wedge dP_i^k + \sum_{k=1}^{n} dQ_i^k \wedge d\ell_i^k = -\sum_{l=1}^{m}\sum_{k=1}^{n} \frac{\partial g^l}{\partial q^k}(Q_i) dQ_i^k \wedge d\Lambda_i^l\,. \tag{3.6}$$

An alternative way to understand this formula is as follows: if the variable λ would be constant then as for unconstrained Hamiltonian systems (see [HNW93, Formula II.16.18]) the expression on the left-hand side of (3.6) would vanish; hence only the terms involving $d\Lambda_i^l$ have to be considered. From the hypotheses (2.3) and the formulas (3.5) and (3.6) we get

$$\sum_{k=1}^{n} dq_1^k \wedge dp_1^k - \sum_{k=1}^{n} dq_0^k \wedge dp_0^k = -h\sum_{i=1}^{s} \widehat{b}_i \sum_{l=1}^{m}\sum_{k=1}^{n} \frac{\partial g^l}{\partial q^k}(Q_i) dQ_i^k \wedge d\Lambda_i^l\,. \tag{3.7}$$

Now, by the use of $g(Q_i)=0$, we have $G(Q_i)dQ_i=0$, i.e.,

$$\sum_{k=1}^{n} \frac{\partial g^l}{\partial q^k}(Q_i)dQ_i^k = 0, \tag{3.8}$$

which finally gives the desired result. □

This result is another motivation to consider the constraints (3.1c) of index 3 and not those of reduced indices (3.1d,e) in Definition 3.1.

RK methods are special cases of PRK methods with coefficients satisfying $\widehat{a}_{ij} = a_{ij}$, $\widehat{b}_i = b_i$, and $\widehat{c}_i = c_i$. In [J93b], [HJ93], [J94], and [J95], the convergence behavior of collocation and RK methods applied to semi-explicit index-3 DAEs in Hessenberg form has been analyzed in detail, confirming the conjecture of [HLR89, p. 86]. Compared to other methods requiring equivalent work, *stiffly accurate* RK methods, i.e., methods which satisfy

$$a_{sj} = b_j \qquad \text{for} \quad j = 1, \ldots, s, \tag{3.9}$$

are tuned to give highly accurate results when applied to DAEs. Unfortunately this later assumption and the symplecticity condition $b_i a_{ij} + b_j a_{ji} - b_i b_j = 0$ lead to $b_s = 0$ and $a_{is} = 0$ for i satisfying $b_i \neq 0$. Therefore we have the following result.

THEOREM 3.2. *No symplectic and stiffly accurate RK schemes exist.*

This negative result is another motivation for the consideration of PRK methods. For RK methods with an invertible RK matrix A, one can also easily show that (3.9) implies that $R(\infty)=0$ where R is the stability function of the method, whereas symplectic schemes must satisfy $|R(\infty)|=1$.

For PRK methods the stiff accuracy condition (3.9) implies that $q_1 = Q_s$ and $g(q_1) = g(Q_s) = 0$. For symplectic PRK methods (see (2.3)) satisfying $b_i \neq 0$ this condition (3.9) implies that

$$\widehat{a}_{js} = 0 \qquad \text{for} \quad j = 1, \ldots, s\,, \tag{3.10}$$

and conversely. Consequently, we now restrict our analysis to PRK methods satisfying (3.10). Under this assumption, Λ_s does not influence the solution of the following nonlinear system originating from (3.3b,c):

$$\begin{aligned} Q_i &= q_0 + h\sum_{j=1}^{s} a_{ij} H_p^T(Q_j, P_j)\,, \\ P_i &= p_0 - h\sum_{j=1}^{s-1} \widehat{a}_{ij}\Big(H_q^T(Q_j, P_j) + G^T(Q_j)\Lambda_j\Big)\,, \qquad 0 = g(Q_i)\,. \end{aligned} \tag{3.11}$$

However, Λ_s enters into the definition of p_1 in (3.3a). It is therefore natural to use this extra freedom by choosing Λ_s such that (q_1, p_1) satisfies (3.1d), i.e., p_1 and Λ_s are the solution of (3.3d)

$$\begin{aligned} p_1 &= p_0 - h\textstyle\sum_{i=1}^{s-1} \widehat{b}_i\Big(H_q^T(Q_i, P_i) + G^T(Q_i)\Lambda_i\Big) - h\widehat{b}_s\Big(H_q^T(Q_s, P_s) + G^T(Q_s)\Lambda_s\Big)\,, \\ 0 &= G(q_1)H_p^T(q_1, p_1)\,. \end{aligned}$$

For constrained Hamiltonian systems this projection onto the manifold (3.1d) does not destroy the symplecticity property shown in Theorem 3.1 (see also [LS94]). The system (3.3d) defines implicitly p_1 and Λ_s in a unique way.

Several definitions of the numerical Lagrange multiplier λ_1 are conceivable. One possibility is to define λ_1 such that (q_1, p_1, λ_1) satisfies (3.1e), i.e., λ_1 is given by

$$\lambda_1 = \left(GH_{pp}^T G^T\right)^{-1} \left(G_q(H_p^T, H_p^T) + GH_{pq}^T H_p^T - GH_{pp}^T H_q^T\right)(q_1, p_1) . \tag{3.3e}$$

However, a very accurate value for λ_1 may be unnecessary. This remark is important if one wants to avoid the computation of extra derivatives such as G_q. For PRK methods which satisfy $c_s = \widehat{c}_s = 1$ a fairly good choice is often given by $\lambda_1 := \Lambda_s$.

Because of the singularity of the matrix $\widehat{A}$ due to (3.10), the nonlinear system (3.11) does not possess a solution in general. This remark applies specifically to the cases where the coefficients (A, b, c) are those of the Radau IIA and Lobatto IIIC methods, and $(\widehat{A}, b, c)$ those of their dual symplectic methods. As it seems obviously necessary to have as many unknowns as equations in (3.11), the only supplementary reasonable assumption to make on the PRK coefficients is

$$a_{1j} = 0 \qquad \text{for} \quad j = 1, \dots, s . \tag{3.12}$$

Thus we get $Q_1 = q_0$, and $g(Q_1) = g(q_0) = 0$ is automatically satisfied if q_0 is consistent. For symplectic PRK methods (see (2.3)) satisfying $b_i \neq 0$, this assumption implies that

$$\widehat{a}_{j1} = b_1 \qquad \text{for} \quad j = 1, \dots, s , \tag{3.13}$$

and conversely. Under the assumptions (3.10) and (3.12) the local existence and uniqueness of the PRK solution (q_1, p_1, λ_1) can be shown provided h is sufficiently small, $\widehat{b}_s \neq 0$, and $A_0 \widehat{A}_0$ is invertible where

$$A_0 = \begin{pmatrix} a_{21} & \dots & a_{2s} \\ \vdots & \ddots & \vdots \\ a_{s1} & \dots & a_{ss} \end{pmatrix}, \qquad \widehat{A}_0 = \begin{pmatrix} \widehat{a}_{11} & \dots & \widehat{a}_{1,s-1} \\ \vdots & \ddots & \vdots \\ \widehat{a}_{s1} & \dots & \widehat{a}_{s,s-1} \end{pmatrix} . \tag{3.14}$$

In this situation, for an efficient solution of the nonlinear system (3.11) (the unknowns are $Q_2, \dots, Q_s, P_1, \dots, P_s, \Lambda_1, \dots, \Lambda_{s-1}$), simplified Newton iterations with the approximate Jacobian matrix

$$\begin{pmatrix} I_{(s-1)n} & -hA_0 \otimes H_{pp}^T(q_0, p_0) & O \\ O & I_{sn} & -h\widehat{A}_0 \otimes G^T(q_0) \\ I_{s-1} \otimes G(q_0) & O & O \end{pmatrix} \tag{3.15}$$

lead to very simple iterations (see also [HLR89, §7]). Only the decomposition of the matrix $(GH_{pp}^T G^T)(q_0, p_0)$ is needed, and at each iteration $s-1$ independent linear systems of dimension m must be solved, this remark being important for a parallel implementation. Due to the invertibility of $A_0 \widehat{A}_0$ and of $(GH_{pp}^T G^T)(q_0, p_0)$, the matrix given in (3.15) is invertible. Another important remark is that high order methods allow the use of larger stepsizes than low order methods. However, the higher the order of the method, the more Newton-type iterations are necessary to preserve this order, and the larger the number s of involved internal stages is required. Therefore a trade-off between a high order and a low number s of internal stages must be made for an efficient implementation.

We mention that a particular class of PRK methods is given by the *half-explicit methods* (*HEMs*) of Ostermann (see [O93]), whose coefficients satisfy

$$\begin{aligned} &a_{j1} = a_{1j} = a_{2j} = 0 , \quad a_{sj} = b_j , \quad \widehat{a}_{s-1,j} = \widehat{b}_j \qquad \text{for} \quad j = 1, \dots, s , \\ &a_{ij} = \widehat{a}_{ij} = 0 \quad \text{if} \quad i \leq j \qquad \text{for} \quad i = 1, \dots, s , \quad j = 1, \dots, s . \end{aligned} \tag{3.16}$$

However, no second order HEM constructed in [O93] is symplectic. The exception is the first order HEM already presented in [O90] and [HLR89, p. 90], which is symplectic if the Hamiltonian is separable.

4. Convergence results. We first introduce the following additional symplifying assumptions which enter into the analysis of PRK methods:

$$C\widehat{C}(Q): \qquad \sum_{j=1}^{s}\sum_{l=1}^{s} a_{ij}\widehat{a}_{jl}c_l^{k-2} = \frac{c_i^k}{k(k-1)} \qquad \text{for} \quad i=1,\dots,s\,, \quad k=2,\dots,Q\,,$$

$$D\widehat{D}(R): \qquad \sum_{i=1}^{s}\sum_{j=1}^{s} b_i c_i^{k-2} a_{ij}\widehat{a}_{jl} = \frac{b_l}{k} - \frac{b_l c_l}{k-1} + \frac{b_l c_l^k}{k(k-1)} \qquad \text{for} \quad l=1,\dots,s\,,$$
$$k=2,\dots,R\,.$$

If the RK coefficients (A, b, c) satisfy $C(q)$ and $D(r)$, and $(\widehat{A}, b, c)$ satisfy $\widehat{C}(\widehat{q})$ and $\widehat{D}(\widehat{r})$ then it can be easily shown that $Q \geq \min(q, \widehat{q}+1)$ and $R \geq \min(\widehat{r}, r+1)$.

From the discussion of the preceding section we only consider here the class of PRK methods with coefficients (A, b, c)–$(\widehat{A}, \widehat{b}, \widehat{c})$ satisfying the following hypotheses:

$$H: \quad \begin{aligned} &a_{1j}=0,\ a_{sj}=b_j,\ \widehat{a}_{js}=0,\ b_j=\widehat{b}_j,\ c_j=\widehat{c}_j \qquad \text{for} \quad j=1,\dots,s\,,\\ &A_0\widehat{A}_0 \text{ is invertible, } b_s\neq 0\,,\\ &B(p),\ C(q),\ D(r),\ \widehat{C}(\widehat{q}),\ \widehat{D}(\widehat{r}),\ C\widehat{C}(Q),\ D\widehat{D}(R)\,. \end{aligned}$$

A detailed analysis of the application of such methods to semi-explicit index-3 DAEs in Hessenberg form is given in [J94]. Here we only state the main results of [J94] related to Hamiltonian systems with holonomic constraints. The analysis of the application of collocation and RK methods to semi-explicit index-3 DAEs in Hessenberg form is given in [HLR89, §6], [J93b], [HJ93], [J94], and [J95]. For PRK methods, slight modifications with regard to RK methods are necessary.

THEOREM 4.1 (local error). (a) *Consider the Hamiltonian system with holonomic constraints* (3.1), *consistent initial values* (q_0, p_0, λ_0) *at* t_0, *and the PRK method* (3.3) *with coefficients* (A,b,c)–$(\widehat{A}, \widehat{b}, \widehat{c})$ *satisfying the hypotheses given by* H. *Then we have*

$$\delta q_h(t_0) = \mathcal{O}(h^{k+1})\,, \quad \delta p_h(t_0) = \mathcal{O}(h^{\ell+1})\,, \quad \delta\lambda_h(t_0) = \mathcal{O}(h^{\ell+1}) \tag{4.1}$$

where

(4.2)

$$k = \min(p, 2q+2, 2\widehat{q}+2, q+r+1, q+\widehat{r}+2, \widehat{q}+r+2, \widehat{q}+\widehat{r}+1, 2Q, Q+\widehat{r}, Q+R)\,,$$

$$\ell = \min(p, 2q+2, 2\widehat{q}+2, q+r+1, q+\widehat{r}+2, \widehat{q}+r+2, \widehat{q}+\widehat{r}+1, 2Q-1, Q+\widehat{r}, Q+R)\,.$$

(b) *If in addition the hypotheses of Theorem* 2.2 *are satisfied then we obtain*

$$\begin{aligned} k &= \min(p, 2q+2, 2r+2, q+r+1, 2Q, Q+q)\,,\\ \ell &= \min(p, 2q+2, 2r+2, q+r+1, 2Q-1)\,. \end{aligned} \tag{4.2′}$$

The *proof* of part (a) which is given in [J94] is rather technical. It makes use of a "rooted-tree-type" theory about the Taylor expansion of the exact and the numerical solutions. The most difficult part is to estimate the local error of the p-component. Part (b) is a direct consequence of part (a) since $\widehat{q}=r$ and $\widehat{r}=q$ (see the first remark after Theorem 2.2). □

THEOREM 4.2 (global error). *Under the same hypotheses stated in Theorem* 4.1 *we have for* $t_n - t_0 = nh \leq$ Const

$$q_n - q(t_n) = \mathcal{O}(h^\ell)\,, \quad p_n - p(t_n) = \mathcal{O}(h^\ell)\,, \quad \lambda_n - \lambda(t_n) = \mathcal{O}(h^\ell), \tag{4.3}$$

where ℓ *is the value given in* (4.2)–(4.2′).

Remark. This theorem remains valid in the case of variable stepsizes with $h = \max_i h_i$.

The *proof* which is given in [J94] is similar to those of [HLR89, Thm. 6.4] and [J95, Thm. 6.1]. It makes use of the following recursion between two neighboring PRK solutions,

$$\begin{aligned} (P_q)_{n+1}\Delta q_{n+1} &= (P_q)_n \Delta q_n + \mathcal{O}\left(h\|(P_q)_n\Delta q_n\| + h\|(P_p)_n \Delta p_n\|\right), \\ (P_p)_{n+1}\Delta p_{n+1} &= (P_p)_n \Delta p_n + \mathcal{O}\left(h\|(P_q)_n\Delta q_n\| + h\|(P_p)_n \Delta p_n\|\right), \end{aligned} \tag{4.4}$$

where P_q and P_p are projectors defined by

$$\begin{aligned} &S := G^T (G H_{pp}^T G^T)^{-1} G\,, \\ &Q_q := H_{pp}^T S\,, \qquad P_q := I - Q_q\,, \qquad Q_p := S H_{pp}^T\,, \qquad P_p := I - Q_p\,. \end{aligned} \tag{4.5}$$

□

The couples of s-stage Lobatto IIIA–IIIB methods (see §2) satisfy the simplifying assumptions $C(s)$, $D(s-2)$, $\widehat{C}(s-2)$, and $\widehat{D}(s)$. Hence $C\widehat{C}(s-1)$ and $D\widehat{D}(s-1)$ must hold. In fact they also satisfy $C\widehat{C}(s)$ and $D\widehat{D}(s)$, and this is the subject of the following lemma.

LEMMA 4.3. *Suppose that* $a_{1j} = 0$ *and* $a_{sj} = b_j$ *hold for* $j = 1, \ldots, s$, *and that the hypotheses of Theorem* 2.2 *are fulfilled with* $p = 2s-2$, $q = s-1$, *and* $r = s-2$. *Then* $C\widehat{C}(s)$ *and* $D\widehat{D}(s)$ *are satisfied.*

Proof of $C\widehat{C}(s)$. Because of $Q \geq s-1$ it is sufficient to show that the coefficients

$$\delta_i := \sum_{j=1}^{s} \sum_{k=1}^{s} a_{ij} \widehat{a}_{jk} c_k^{s-2} - \frac{c_i^s}{s(s-1)} \tag{4.6}$$

vanish for $i = 1, \ldots, s$. From $a_{1j} = 0$ and $c_1 = 0$ we have $\delta_1 = 0$. Using $a_{sj} = b_j$ and $c_s = 1$ we get

$$\begin{aligned} \delta_s &= \sum_{k=1}^{s} \sum_{j=1}^{s} b_j \widehat{a}_{jk} c_k^{s-2} - \frac{1}{s(s-1)} \overset{\widehat{D}(1)}{=} \sum_{k=1}^{s} b_k (1-c_k) c_k^{s-2} - \frac{1}{s(s-1)} \\ &\overset{B(s)}{=} \frac{1}{s-1} - \frac{1}{s} - \frac{1}{s(s-1)} = 0\,. \end{aligned} \tag{4.7}$$

We will next show that the sums

$$S_m := \sum_{i=1}^{s} b_i c_i^{m-1} \delta_i \tag{4.8}$$

vanish for $m = 1, \ldots, s-2$. This will give the desired result $\delta_i = 0$ for $i = 2, \ldots, s-1$. By the symplecticity condition $a_{ij} = b_j(1 - \widehat{a}_{ji}/b_i)$ we get $S_m = A_m - B_m - C_m$ where

$$A_m = \sum_{i=1}^{s} b_i c_i^{m-1} \sum_{k=1}^{s} \sum_{j=1}^{s} b_j \widehat{a}_{jk} c_k^{s-2}\,, \tag{4.9}$$

$$B_m = \sum_{j=1}^{s} b_j \sum_{i=1}^{s} \widehat{a}_{ji} c_i^{m-1} \sum_{k=1}^{s} \widehat{a}_{jk} c_k^{s-2}\,, \tag{4.10}$$

$$C_m = \frac{1}{s(s-1)} \sum_{i=1}^{s} b_i c_i^{m+s-1} . \tag{4.11}$$

Each term can be computed separately:

$$A_m \overset{\widehat{D}(1)}{=} \sum_{i=1}^{s} b_i c_i^{m-1} \sum_{k=1}^{s} b_k (1-c_k) c_k^{s-2} \overset{B(s)}{=} \frac{1}{m} \left(\frac{1}{s-1} - \frac{1}{s} \right) = \frac{1}{ms(s-1)} , \tag{4.12}$$

$$B_m \overset{\widehat{C}(s-2)}{=} \frac{1}{m} \sum_{k=1}^{s} \sum_{j=1}^{s} b_j c_j^m \widehat{a}_{jk} c_k^{s-2} \overset{\widehat{D}(s-1)}{=} \frac{1}{m(m+1)} \sum_{k=1}^{s} b_k (1-c_k^{m+1}) c_k^{s-2} \tag{4.13}$$

$$\overset{B(2s-2)}{=} \frac{1}{m(m+1)} \left(\frac{1}{s-1} - \frac{1}{m+s} \right) = \frac{1}{m(m+s)(s-1)} ,$$

$$C_m \overset{B(2s-2)}{=} \frac{1}{(m+s)s(s-1)} . \tag{4.14}$$

From these results we easily get $S_m = 0$.

Proof of $D\widehat{D}(s)$. Because of $R \geq s-1$ it is sufficient to show that the coefficients

$$\mu_k := \sum_{i=1}^{s} \sum_{j=1}^{s} b_i c_i^{s-2} a_{ij} \widehat{a}_{jk} - b_k \left(\frac{1}{s} - \frac{c_k}{s-1} + \frac{c_k^s}{s(s-1)} \right) \tag{4.15}$$

vanish for $k = 1, \ldots, s$. From $\widehat{a}_{js} = 0$ (see (3.7)) and $c_s = 1$ we have $\mu_s = 0$. Using $\widehat{a}_{j1} = b_1$ (see (3.10)) and $c_1 = 0$ we get

$$\mu_1 = b_1 \left(\sum_{i=1}^{s} b_i c_i^{s-1} - \frac{1}{s} \right) \overset{B(s)}{=} 0 . \tag{4.16}$$

We will next show that the sums

$$T_m := \sum_{k=1}^{s} \mu_k c_k^{m-1} \tag{4.17}$$

vanish for $m = 1, \ldots, s-2$. This will give the desired result $\mu_k = 0$ for $k = 2, \ldots, s-1$. By the symplecticity condition $\widehat{a}_{jk} = b_k (1 - a_{kj}/b_j)$ we get $T_m = D_m - E_m - F_m$ where

$$D_m = \sum_{i=1}^{s} b_i c_i^{s-2} \sum_{j=1}^{s} a_{ij} \sum_{k=1}^{s} b_k c_k^{m-1} , \tag{4.18}$$

$$E_m = \sum_{i=1}^{s} b_i c_i^{s-2} \sum_{j=1}^{s} \frac{a_{ij}}{b_j} \sum_{k=1}^{s} b_k c_k^{m-1} a_{kj} , \tag{4.19}$$

$$F_m = \frac{1}{s} \sum_{k=1}^{s} b_k c_k^{m-1} - \frac{1}{s-1} \sum_{k=1}^{s} b_k c_k^m + \frac{1}{s(s-1)} \sum_{k=1}^{s} b_k c_k^{m+s-1} . \tag{4.20}$$

Each term can be computed separately:

$$D_m \overset{C(1)}{=} \sum_{i=1}^{s} b_i c_i^{s-1} \sum_{k=1}^{s} b_k c_k^{m-1} \overset{B(s)}{=} \frac{1}{ms}, \tag{4.21}$$

$$E_m \overset{D(s-2)}{=} \frac{1}{m} \sum_{i=1}^{s} b_i c_i^{s-2} \sum_{j=1}^{s} a_{ij}(1-c_j^m) \tag{4.22}$$

$$\overset{C(s-1)}{=} \frac{1}{m} \sum_{i=1}^{s} b_i c_i^{s-1} - \frac{1}{m(m+1)} \sum_{i=1}^{s} b_i c_i^{m+s-1}$$

$$\overset{B(2s-2)}{=} \frac{1}{ms} - \frac{1}{m(m+1)(m+s)},$$

$$F_m \overset{B(2s-2)}{=} \frac{1}{ms} - \frac{1}{(m+1)(s-1)} + \frac{1}{s(s-1)(m+s)}. \tag{4.23}$$

From these results we easily get $T_m = 0$. □

From Theorem 4.1, Theorem 4.2, and Lemma 4.3 we have the following convergence result.

COROLLARY 4.4. *For the symplectic couples of s-stage Lobatto IIIA–IIIB methods applied to the constrained Hamiltonian system* (3.1) *(see* (3.3)*) with consistent initial values* (q_0, p_0, λ_0) *at* t_0*, the global error satisfies, for* $t_n - t_0 = nh \leq$ Const,

$$q_n - q(t_n) = \mathcal{O}(h^{2s-2}), \quad p_n - p(t_n) = \mathcal{O}(h^{2s-2}), \quad \lambda_n - \lambda(t_n) = \mathcal{O}(h^{2s-2}). \tag{4.24}$$

Proof. The symplecticity of Lobatto IIIA–IIIB methods has been proved in [Sun92]. These methods satisfy the hypotheses given by Theorem 2.2 and H with $p = 2s-2$, $q = \widehat{r} = s$, $r = \widehat{q} = s-2$, and $Q = R = s$. The invertibility of the matrix $A_0 \widehat{A}_0$ simply follows from $C\widehat{C}(s)$. □

For separable Hamiltonian systems the 2-stage Lobatto IIIA–IIIB method is half-explicit and is equivalent to the *Rattle algorithm* proposed in [An83] (see also [LS94]).

Because of the presence of the "explicit" stage P_s in (2.1) and (3.3) for symplectic PRK methods satisfying $a_{sj} = b_j$, the Lobatto IIIA–IIIB methods are not appropriate when solving stiff Hamiltonian systems, e.g., Hamiltonian systems containing a strong potential of the form

$$\frac{1}{\varepsilon^2} V(q), \qquad 0 < \varepsilon \ll 1. \tag{4.25}$$

This has been numerically observed when trying to solve the stiff spring pendulum equations (see [Lu93] and [HLR89, pp. 10–12]) with Lobatto IIIA–IIIB methods.

For the long-time integration of Hamiltonian systems, a constant-stepsize application of symplectic methods performs generally better than variable-stepsize algorithms if the time-scale does not vary greatly along the solution (see [CSS92]). The reason lies in a "backward analysis" argument (see [SS92] and [H94]). For constant stepsizes and symplectic methods, the numerical solution can be interpreted as the exact solution of a nearby perturbed Hamiltonian system. It is likely that this result can be extended to Hamiltonian systems with holonomic constraints. We also point out that the construction of an embedded PRK scheme is not crucial for a constant-stepsize implementation if an approximation to the error of the method is not needed.

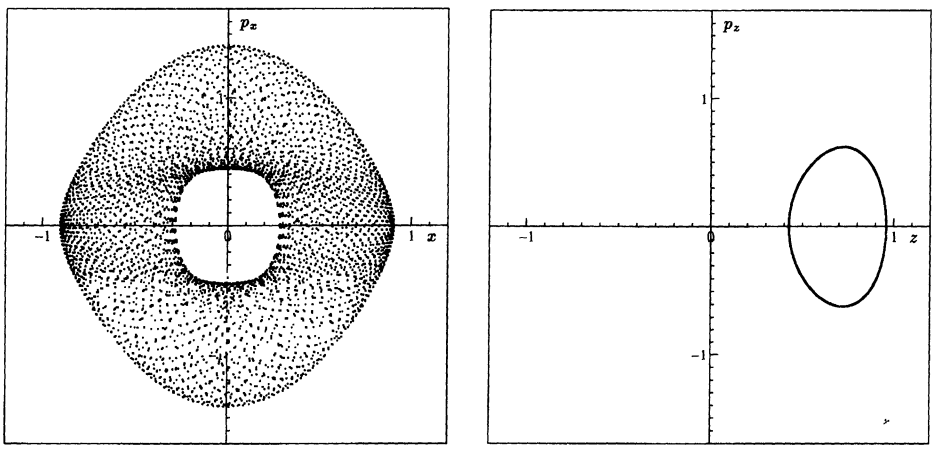

FIG. 5.1. *The phase portraits* (x, p_x) *and* (z, p_z) *of the 3-stage Lobatto IIIA–IIIB method applied to Example 1 with stepsize* $h = 0.12$.

5. Numerical experiments. We first notice that for the solution of the nonlinear system (3.11), the s-stage Lobatto IIIA–IIIB method requires a computational work approximately equivalent to that arising for the $(s-1)$-stage RK methods Radau IIA and Gauss. Hence these methods are comparable.

Example 1. We consider the motion of a particle of mass m and electric charge e, moving on a sphere of radius R under the action of forces due to an electric field $(0, 0, E)^T$ and to a magnetic field $(0, 0, B)^T$ (see [Cho92, Problem 7.16]). We use the cartesian coordinates $q = (x, y, z)^T$ for the description of the position of the particle. The holonomic constraint is expressed by

$$0 = \sqrt{x^2+y^2+z^2} - R \,. \tag{5.1}$$

Denoting the generalized momenta by $p = (p_x, p_y, p_z)^T$, the Hamiltonian of the system is given by

$$H(q, p) = \frac{1}{2m}\Big((p_x+m\omega y)^2+(p_y-m\omega x)^2+p_z^2\Big) - eEz \tag{5.2}$$

where $\omega = eB/(2mc)$ and c is the speed of light. This is an example of a nonseparable Hamiltonian.

We have applied 5000 steps of the 3-stage Lobatto IIIA–IIIB method of order 4 with stepsize $h = 0.12$,

$$m = 1, \qquad \omega = 1, \qquad R = 1, \qquad eE = 1 \,, \tag{5.3}$$

and consistent initial values

$$\begin{aligned} &x(0) = 0.2, \qquad y(0) = 0.2, \qquad z(0) = \sqrt{0.92}, \\ &p_x(0) = 1, \qquad p_y(0) = -1, \qquad p_z(0) = 0 \,. \end{aligned} \tag{5.4}$$

We have plotted in Fig. 5.1 the phase portraits (x, p_x) and (z, p_z).

In Fig. 5.2 we have drawn the first 500 steps of the numerical Hamiltonian, whose value for the exact solution is $H = 1.2^2 - \sqrt{0.92} \approx 0.4808336955$. The remaining 4500 steps show

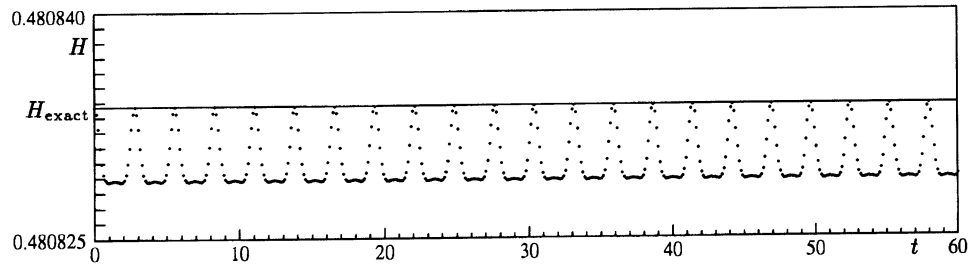

FIG. 5.2. *The numerical Hamiltonian of the* 3*-stage Lobatto IIIA–IIIB method applied to Example* 1 *with stepsize* $h=0.12$.

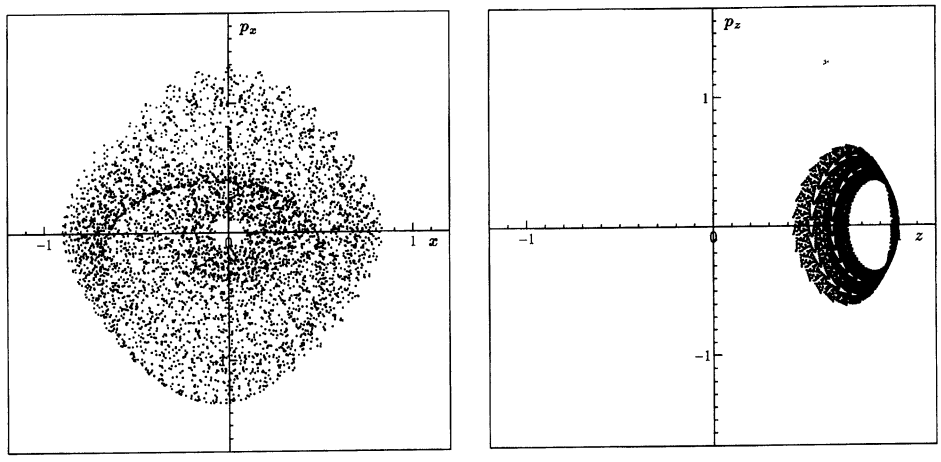

FIG. 5.3. *The phase portraits* (x, p_x) *and* (z, p_z) *of the* 2*-stage Radau IIA method applied to Example* 1 *with stepsize* $h=0.12$.

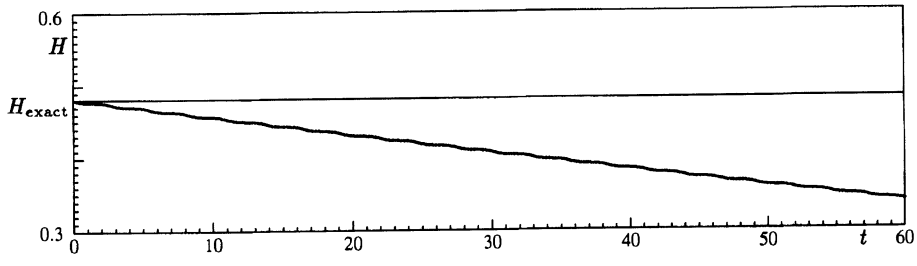

FIG. 5.4. *The numerical Hamiltonian of the* 2*-stage Radau IIA method applied to Example* 1 *with stepsize* $h=0.12$.

the same periodic behavior. If the scale of Fig. 5.4 would be used here, then the numerical Hamiltonian would appear nearly equal to the exact value.

As a comparison we have applied the 2-stage Radau IIA method with the same stepsize $h=0.12$. The numerical results are given in Figs. 5.3 and 5.4.

Since the numerical solution of the Radau IIA method does not satisfy all underlying constraints, we have also applied this method with projections onto these constraints after every step (see [AP91], [J94], and [J95] for similar ideas). Although the theoretical order of convergence is improved compared to the unprojected method, the numerical results did

not exhibit any visible difference with regards to Figs. 5.3 and 5.4. For that reason the corresponding figures are not plotted.

We observe that for the Lobatto IIIA–IIIB method the numerical Hamiltonian remains in tolerable bounds, but it drifts away from the exact value (roughly linearly with time) for the unprojected and projected Radau IIA methods. This is a demonstration of the different behaviors of symplectic and nonsymplectic integrators.

Example 2. The double pendulum. We use the cartesian coordinates $q_1 = (x_1, z_1)^T$, $q_2 = (x_2, z_2)^T$ for the description of the position of each pendulum. The two holonomic constraints on the lengths ℓ_1 and ℓ_2 of the two pendula are

$$0 = \sqrt{x_1^2+z_1^2}-\ell_1, \qquad 0 = \sqrt{(x_2-x_1)^2+(z_2-z_1)^2}-\ell_2\ . \tag{5.5}$$

Denoting the generalized momenta by $p_1=(p_{x_1}, p_{z_1})^T$ and $p_2=(p_{x_2}, p_{z_2})^T$, the Hamiltonian of the system is given by

$$H(q, p) = \frac{1}{2m_1}\left(p_{x_1}^2+p_{z_1}^2\right)+\frac{1}{2m_2}\left(p_{x_2}^2+p_{z_2}^2\right)+m_1gz_1+m_2gz_2 \tag{5.6}$$

and is separable.

We have applied 5000 steps of the 3-stage Lobatto IIIA–IIIB method of order 4 with stepsize $h=0.12$,

$$m_1 = 1 = m_2, \qquad \ell_1 = 1 = \ell_2, \qquad g = 1\ , \tag{5.7}$$

and consistent initial values

$$\begin{aligned} &x_1(0) = 0.5, \qquad z_1(0) = -\sqrt{0.75}, \qquad x_2(0) = 0, \qquad z_2(0) = -2\sqrt{0.75}, \\ &p_{x_1}(0) = 0, \qquad p_{z_1}(0) = 0\ , \qquad p_{x_2}(0) = 0, \qquad p_{z_2}(0) = 0\ . \end{aligned} \tag{5.8}$$

We have plotted in Fig. 5.5 the phase portraits (x_1, p_{x_1}), (z_2, p_{z_2}), and in Fig. 5.6 the first 500 steps of the numerical Hamiltonian whose value for the exact solution is $H = -3\sqrt{0.75} \approx -2.5980762113$.

As a comparison we have applied the projected 2-stage Gauss method and the unprojected and projected 2-stage Radau IIA methods to this problem with the same stepsize $h = 0.12$. Their numerical Hamiltonian is plotted in Fig. 5.7. We point out that the unprojected 2-stage Gauss method generally diverges when applied to Hamiltonian systems with holonomic constraints (see [HJ93]). This has been numerically observed for this problem. Although the unprojected Gauss methods are symplectic, the projected Gauss methods are not, and we clearly see here that the numerical Hamiltonian drifts off the exact value. However, this drift is less drastic here than for the unprojected Radau IIA method, which in turn is less severe than for the projected Radau IIA method.

In Fig. 5.8, as a last experiment, the global errors at $t = 5$ of the four above-mentioned methods have been plotted as functions of h. Since we have used logarithmic scales, the curves appear as straight lines of slope k whenever the leading term of the error is $\mathcal{O}(h^k)$. This behavior is indicated in the figures.

The order of convergence of the projected s-stage Gauss method is s and that of the projected s-stage Radau IIA method is $2s-1$ (see [J93b], [J94], and [J95]). For the unprojected s-stage Radau IIA method the order of convergence is $2s-1$ for the q-component, s for the p-component, and s–1 for the λ-component. The predicted orders are confirmed in Fig. 5.8,

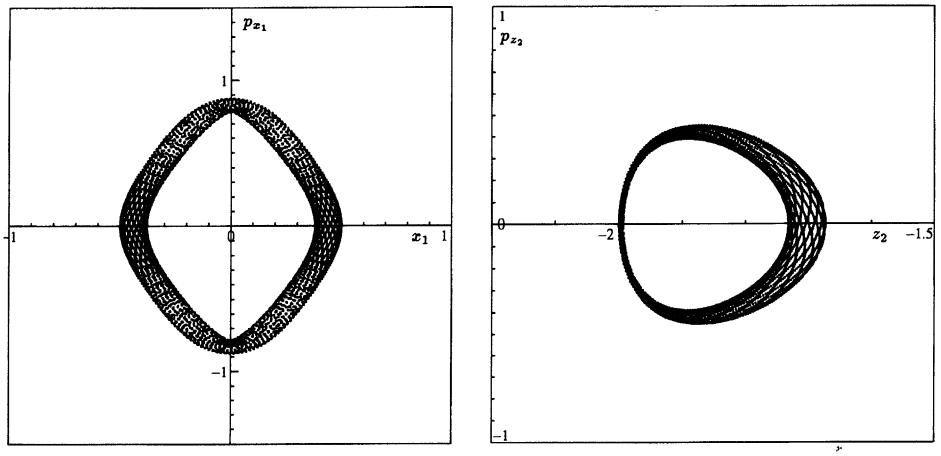

FIG. 5.5. *The phase portraits* (x_1, p_{x_1}) *and* (z_2, p_{z_2}) *of the 3-stage Lobatto IIIA–IIIB method applied to Example 2 with stepsize* $h = 0.12$.

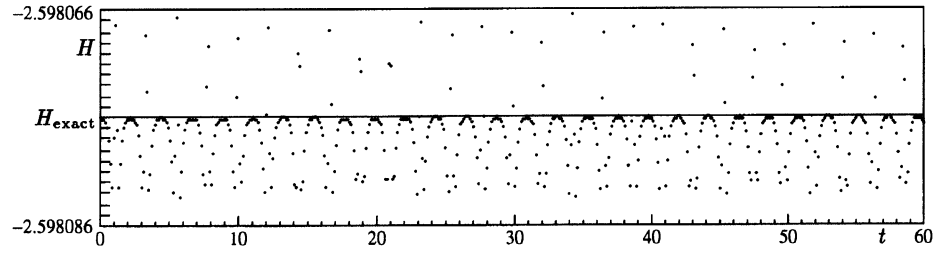

FIG. 5.6. *The numerical Hamiltonian of the 3-stage Lobatto IIIA–IIIB method applied to Example 2 with stepsize* $h = 0.12$.

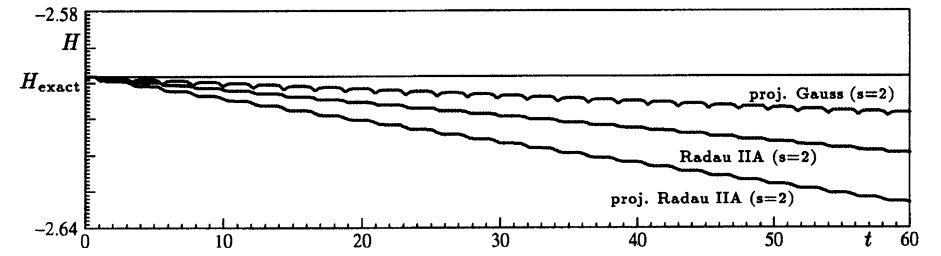

FIG. 5.7. *The numerical Hamiltonians of the projected 2-stage Gauss method and of the unprojected and projected 2-stage Radau IIA methods applied to Example 2 with stepsize* $h = 0.12$.

and this clearly shows the superiority of the Lobatto IIIA–IIIB schemes, also in terms of accuracy.

Acknowledgments. I am grateful to E. Hairer for his remarks during the preparation of this paper. I express my gratitude to G. Wanner for his comments on a preliminary version of this article. I wish to thank S. F. Bernatchez for her careful reading of the English part of the manuscript. Special thanks go to the unknown referees for their constructive criticisms.

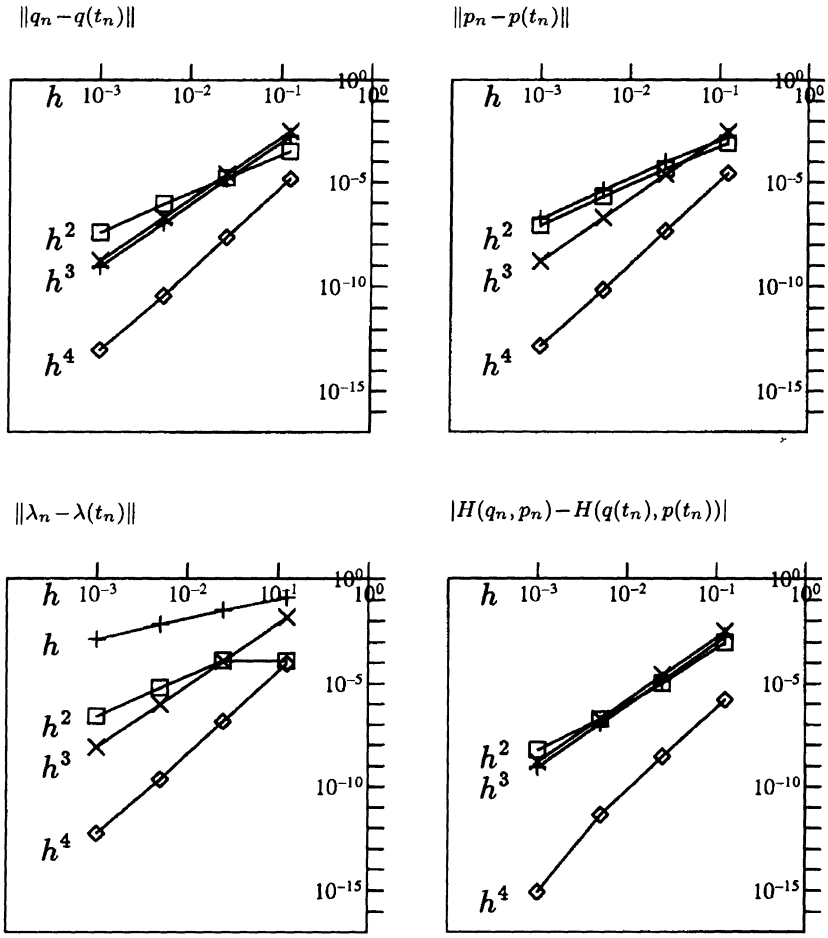

FIG. 5.8. *Global errors at $t = 5$ of four methods applied to Example 2 (projected 2-stage Gauss: □; projected 2-stage Radau IIA: ×; 2-stage Radau IIA: +; 3-stage Lobatto IIIA–IIIB: ◇).*

REFERENCES

[ASS93] L. ABIA AND J. M. SANZ-SERNA, *Partitioned Runge-Kutta methods for separable Hamiltonian problems*, Math. Comput., 60 (1993), pp. 617–634.

[An83] H. C. ANDERSEN, *Rattle: A velocity version of the Shake algorithm for molecular dynamics calculations*, J. Comput. Phys., 52 (1983), pp. 24–34.

[Ar89] V. I. ARNOLD, *Mathematical Methods of Classical Mechanics*, Second ed., Graduate Texts in Mathematics, Vol. 60, Springer-Verlag, New York, 1989.

[A85] U. M. ASCHER, *Two families of symmetric difference schemes for singular perturbation problems*, in Numerical Boundary Values ODEs, U. M. Ascher and R. D. Russel, eds., Proceedings of an International Workshop, Vancouver, Canada, July 10–13, 1984, Progress in Scientific Computing 5, Birkhäuser, Boston, 1985, pp. 173–191.

[AMR88] U. M. ASCHER, R. M. M. MATTHEIJ, AND R. D. RUSSEL, *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*, Prentice-Hall, Englewood Cliffs, NJ, 1988.

[AP91] U. M. ASCHER AND L. R. PETZOLD, *Projected implicit Runge–Kutta methods for differential-algebraic equations*, SIAM J. Numer. Anal., 28 (1991), pp. 1097–1120.

[AP92] ———, *Projected collocation methods for higher-order higher-index differential-algebraic equations*, J. Comput. Appl. Math., 43 (1992), pp. 243–259.

[AP93] U. M. ASCHER AND L. R. PETZOLD, *Stability of computational methods for constrained dynamics systems*, SIAM J. Sci. Comput., 14 (1993), pp. 95–120.

[BCP89] K. E. BRENAN, S. L. CAMPBELL, AND L. R. PETZOLD, *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*, North-Holland, New York, 1989.

[Bu87] J. C. BUTCHER, *The Numerical Analysis of Ordinary Differential Equations. Runge-Kutta and General Linear Methods*, John Wiley and Sons, Chichester, 1987.

[Cha90] R. P. K. CHAN, *On symmetric Runge-Kutta methods of high order*, Computing, 45 (1990), pp. 301–309.

[Cho92] P. CHOQUARD, *Mécanique Analytique*, Vol. 1, Cahiers mathématiques de l'Ecole Polytechnique Fédérale de Lausanne, Presses Polytechniques et Universitaires Romandes, Lausanne, Switzerland, 1992.

[CSS92] M. P. CALVO AND J. M. SANZ-SERNA, *Variable steps for symplectic integrators*, in Numerical Analysis 1991, D. F. Griffiths and G. A. Watson, eds., Pitman Research Notes in Math. Series 260, Longman, Essex, England, 1992, pp. 34–48.

[CSS94] ———, *Canonical B-series*, Numer. Math., 67 (1994), pp. 161–175.

[FL89] C. FÜHRER AND B. J. LEIMKUHLER, *Formulation and numerical solution of the equations of constrained mechanical motion*, Research Report, DFVLR-FB 89-08, Institut für Dynamik der Flugsysteme, Oberpaffenhofen, Germany, 1989.

[H94] E. HAIRER, *Backward analysis of numerical integrators and symplectic methods*, in Scientific Computation and Differential Equations, K. Burrage et al., eds., Proceedings of the SCADE'93 conference, Auckland, New Zealand, January 1993, Annals of Numerical Mathematics 1, J. C. Baltzer, Amsterdam, 1994, pp. 107–132.

[HJ93] E. HAIRER AND L. JAY, *Implicit Runge-Kutta methods for higher index differential-algebraic systems*, in Contributions in Numerical Mathematics, WSSIAA 2, World Scientific, Singapore, 1993, pp. 213–224.

[HLR89] E. HAIRER, CH. LUBICH, AND M. ROCHE, *The Numerical Solution of Differential-Algebraic Systems by Runge-Kutta Methods*, Lecture Notes in Math. 1409, Springer-Verlag, Berlin, New York, 1989.

[HNW93] E. HAIRER, S. P. NØRSETT, AND G. WANNER, *Solving Ordinary Differential Equations* I. *Nonstiff Problems*, Second revised ed., Computational Mathematics, Vol. 8, Springer-Verlag, Berlin, 1993.

[HW91] E. HAIRER AND G. WANNER, *Solving Ordinary Differential Equations* II. *Stiff and Differential-Algebraic Problems*, Computational Mathematics, Vol. 14, Springer-Verlag, Berlin, 1991.

[J93a] L. JAY, *Convergence of a class of Runge-Kutta methods for differential-algebraic systems of index* 2, BIT, 33 (1993), pp. 137–150.

[J93b] ———, *Collocation methods for differential-algebraic equations of index* 3, Numer. Math., 65 (1993), pp. 407–421.

[J94] ———, *Runge-Kutta type methods for index three differential-algebraic equations with applications to Hamiltonian systems*, Ph.D. thesis, Dept. of Math., University of Geneva, Switzerland, 1994.

[J95] ———, *Convergence of Runge-Kutta methods for differential-algebraic systems of index* 3, Appl. Numer. Math., 17 (1995), pp. 97–118.

[La88] F. M. LASAGNI, *Canonical Runge-Kutta methods*, ZAMP, 39 (1988), pp. 952–953.

[LR94] B. J. LEIMKUHLER AND S. REICH, *Symplectic integration of constrained Hamiltonian systems*, Math. Comput., 63 (1994), pp. 589–606.

[LS94] B. J. LEIMKUHLER AND R. D. SKEEL, *Symplectic numerical integrators in constrained Hamiltonian systems*, J. Comput. Phys., 112 (1994), pp. 117–125.

[Lu93] CH. LUBICH, *Integration of stiff mechanical systems by Runge-Kutta methods*, ZAMP, 44 (1993), pp. 1022–1053.

[MK92] R. S. MACKAY, *Some aspects of the dynamics and numerics of Hamiltonian systems*, in The Dynamics of Numerics and the Numerics of Dynamics, D. S. Broomhead and A. Iserles, eds., Clarendon Press, Oxford, 1992, pp. 137–193.

[O90] A. OSTERMANN, *A half-explicit extrapolation method for differential-algebraic systems of index* 3, IMA J. Numer. Anal., 10 (1990), pp. 171–180.

[O93] ———, *A class of half-explicit Runge-Kutta methods for differential-algebraic systems of index* 3, Appl. Numer. Math., 13 (1993), pp. 165–179.

[PP92] L. R. PETZOLD AND F. A. POTRA, *ODAE methods for the numerical solution of Euler-Lagrange equations*, Appl. Numer. Math., 10 (1992), pp. 397–413.

[R93] S. REICH, *Symplectic integration of constrained Hamiltonian systems by Runge-Kutta methods*, Report, Institut für Angewandte Analysis und Stochastik, Berlin, 1993.

[SBO93] R. D. SKEEL, J. J. BIESIADECKI, AND D. OKUNBOR, *Symplectic integration for macromolecular dynamics*, in Proceedings of the International Conference on Computation of Differential Equations and Dynamical Systems, Beijing, China, Sept. 1992, K. Feng and Z.-C. Shi, eds., Series on Appl. Math. 4, World Scientific, Singapore, 1993.

[SS88] J. M. SANZ-SERNA, *Runge-Kutta schemes for Hamiltonian systems*, BIT, 28 (1988), pp. 877–883.

[SS92] J. M. SANZ-SERNA, *Symplectic integrators for Hamiltonian problems: An overview*, Acta Numerica, 1 (1992), pp. 243–286.

[Sun92] SUN GENG, *Symplectic partitioned Runge-Kutta methods*, J. Comput. Math., to appear.

[Sun93] ———, *Construction of high order symplectic Runge-Kutta methods*, J. Comput. Math., 11 (1993), pp. 250–260.

[Sur89] Y. B. SURIS, *On canonicality of mappings resulting from Runge-Kutta-type methods during the integration of the system* $\ddot{x} = -\partial U/\partial x$, Zh. Vychisl. Mat. i Mat. Fiz., 29 (1989), pp. 202–211. (In Russian.) Also in U.S.S.R. Comput. Maths. Phys., 29 (1989), pp. 138–144.

[Y90] H. YOSHIDA, *Construction of higher order symplectic integrators*, Phys. Lett. A, 150 (1990), pp. 262–268.