Symplectic Integration of Constrained Hamiltonian Systems by Runge-Kutta Methods

by
Sebastian Reich

Technical Report 93-13
April 1993

Department of Computer Science
University of British Columbia
Rm 333 - 6356 Agricultural Road
Vancouver, B.C.
CANADA V6T 1Z2

Telephone: (604) 822-3061
Fax: (604) 822-5485

SYMPLECTIC INTEGRATION OF CONSTRAINED HAMILTONIAN SYSTEMS BY RUNGE–KUTTA METHODS

SEBASTIAN REICH
INSTITUT FÜR ANGEWANDTE ANALYSIS UND STOCHASTIK
MOHRENSTRAßE 39
BERLIN, O–1086, GERMANY
E–MAIL: NA.REICH@NA-NET.ORNL.GOV *

Abstract. Recent work reported in the literature suggests that for the long–term integration of Hamiltonian dynamical systems one should use methods that preserve the symplectic structure of the flow. In this paper we investigate the symplecticity of numerical integrators for constrained Hamiltonian systems. In the first part of the paper we show that those implicit Runge–Kutta methods which result in symplectic integrators for unconstrained Hamiltonian systems can be directly applied to constrained Hamiltonian systems. The resulting discretization scheme is symplectic but does not, in general, preserve the constraints. In the second part of the paper we discuss partitioned Runge–Kutta methods. Again it turns out that those partitioned Runge–Kutta methods which are symplectic for unconstrained systems can be applied to constrained Hamiltonian systems. We show that, in contrast to implicit Runge–Kutta methods, the class of symplectic partitioned Runge–Kutta methods includes methods that also preserve the constraints. In the third part of the paper we discuss constrained Hamiltonian systems with separable Hamiltonian from a Lie algebraic point of view. This approach not only provides a different approach to the numercial integration of Hamiltonian systems but also allows for a straighforward backward error analysis.

Key words. Hamiltonian systems, differential–algebraic equations, canonical discretization schemes, Lie groups

AMS(MOS) subject classifications. 65L05, 65L20

* Current address: Department of Computer Science at the University of British Columbia, Vancouver, B.C. V6T 1Z2. This work was partially supported under NSERC Canada Grant OGP0004306.

1. Introduction. There has been much recent interest in the numerical integration of Hamiltonian systems

$$
\begin{aligned}
q' &= +\nabla_p H(q,p) \\
p' &= -\nabla_q H(q,p)
\end{aligned} \tag{1}
$$

where $q,p \in \mathbf{R}^n$ and $H : \mathbf{R}^{2n} \to \mathbf{R}$ is sufficiently smooth. The system (1) can be accurately solved over long time intervals by an implicit canonical discretization scheme [22] which maintains the symplectic structure of the flow. A natural question is what happens when (1) is constrained by algebraic equations on q and/or p. In this paper, we restrict ourselves to the case of holonomic contraints, in which case, starting from a Lagrangian variational principle, one would arrive at a system of differential–algebraic equations (DAEs) of the form:

$$
\begin{aligned}
q' &= +\nabla_p H(q,p) \\
p' &= -\nabla_q H(q,p) - G(q)^t\lambda \\
0 &= g(q)
\end{aligned} \tag{2}
$$

where $g : \mathbf{R}^n \to \mathbf{R}^m$, $G(q) = g_q(q) \in \mathbf{R}^{m\times n}$, and $\lambda \in \mathbf{R}^m$. If

$$G(q)H_{pp}(q,p)G(q)^t \tag{3}$$

is invertible, then (2) defines implicitly a Hamiltonian vector field on the $(2n-2m)$–dimensional manifold

$$\mathcal{M} = \{(q,p) : g(q) = 0,\ G(q)\nabla_p H(q,p) = 0\} \tag{4}$$

The flow of a Hamiltonian vector field possesses an important symplectic structure [2]. Much recent research has gone into developing symplectic numerical discretization schemes for unconstrained systems (1) that inherit the symplectic structure of the original system. It has been observed in numerical experiments [20] that symplectic methods with fixed stepsize possess better long–term stability properties than nonsymplectic methods or symplectic methods with varying stepsize.

The symplectic integration of the constrained equations (2) was treated in [16] via symplectic parametrization of the constraint manifold and by methods based on Dirac's theory on weak invariants. These methods lead to unconstrained Hamiltonian systems of type (1) which can be handled by direct application of symplectic methods. In [17], it was shown that in the case of a separable Hamiltonian function

$$H(q,p) = \frac{p^t M^{-1} p}{2} + V(q) \tag{5}$$

where M is a symmetric, positive definite matrix, the constrained system (2) can be integrated directly by proper modifications [1] of the Verlet scheme. This discretization results in a symplectic method which also preserves the constraints. Note that separable Hamiltonians occur, e.g., in the context of multibody systems and molecular dynamics [25]. A non–separable Hamiltonian is obtained, e.g., whenever a particle is moving under the influence of an electro–magnetic field [9].

In the first part of this paper, we consider direct symplectic numerical discretization of constrained Hamiltonian systems via implicit Runge–Kutta (IRK) methods. We show that those IRK methods which are symplectic for unconstrained systems

(1) lead also to symplectic integrators for constrained Hamiltonian systems with arbitrary Hamiltonian H. However, it turns out that those methods do not preserve the constraint manifold $\mathcal{M}$. Moreover, and at times more importantly, a reduction in the method's order may result.

In the second part, we discuss the discretization of constrained Hamiltonian systems by means of partitioned Runge–Kutta (PRK) methods. Again it turns out that those PRK methods which are symplectic for unconstrained systems (1) can be applied to constrained Hamiltonian systems. Furthermore, we show that there exist symplectic PRK methods that preserve the constraint manifold $\mathcal{M}$. Based on a first order symplectic PRK method, we construct methods of second and higher order following an idea given in [27] for unconstrained systems.

In the third part, we discuss constrained Hamiltonian systems with separable Hamiltonian (5). The generalization of a family of PRK methods introduced first by Ruth in [21] to constrained systems is discussed from a Lie algebraic point of view [8],[19],[27]. Besides providing a different view at the symplectic integration of constrained Hamiltonian systems, this approach allows for a straightforward backward error analysis of symplectic schemes. We show in particular that the numerical solutions can formally be interpreted as the exact solutions of a certain perturbed Hamiltonian system evaluated at discrete time points.

2. Constrained Hamiltonain Systems and Symplectic Structure. For notational convenience we introduce the following notation. Let $x = (q,p) \in \mathbf{R}^{2n}$ and define the mapping $\phi : \mathbf{R}^{2n} \to \mathbf{R}^{2m}$ by

$$\phi(x) = \begin{pmatrix} \phi^1(x) \\ \phi^2(x) \end{pmatrix} = \begin{pmatrix} g(q) \\ G(q)\nabla_p H(q,p) \end{pmatrix}$$

Let $J \in \mathbf{R}^{2n\times 2n}$ be the skew–symmetric matrix

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

Then the dynamics of the constrained Hamiltonian system (2) can be described equivalently by the formulation

$$\begin{aligned} x' &= J\nabla H(x) + J\Phi(x)^t\lambda \\ 0 &= \phi(x) \end{aligned} \tag{6}$$

where $\Phi(x) = \phi_x(x)$ and $\lambda \in \mathbf{R}^{2m}$. Note that the invertibility of (3) implies that $\Phi(x)J\Phi(x)^t$ has full rank too. Thus, in contrast to the index three formulation (2), (6) is a DAE of index two [6]. It is well–known [6], [15] that this fact is significant for the numerical treatment of the problem.

To see that (6) is indeed equivalent to (2) rewrite (6) as

$$\begin{aligned} x' &= J\nabla H(x) + J\Phi^1(x)^t\lambda^1 + J\Phi^2(x)^t\lambda^2 \\ 0 &= \phi^1(x) \\ 0 &= \phi^2(x) \end{aligned} \tag{7}$$

and (2) as

$$\begin{aligned} x' &= J\nabla H(x) + J\Phi^1(x)^t\lambda \\ 0 &= \phi^1(x) \end{aligned}$$

where $\Phi^1(x) = \phi^1_x(x)$, $\Phi^2(x) = \phi^2_x(x)$, and $\lambda^1 \in \mathbf{R}^m$, $\lambda^2 \in \mathbf{R}^m$, $\lambda \in \mathbf{R}^m$. Differentiating the constraints we obtain $\lambda^2 = 0$ and $\lambda^1 = \lambda$. Here one has to make use of the fact that $\Phi^i(x) J \Phi^i(x)^t = 0$ for $i = 1, 2$ and $\phi^2(x) = \Phi^1(x) J \nabla H(x)$.

REMARK 2.1. *(6) represents only one possible reformulation of the index three DAE (2) as an index two problem [5], [10]. However, the results obtained in this paper on symplectic integrators apply only to the index two formulation (6).*

In [16], the following method for the numerical treatment of (2) was proposed. In terms of the formulation (6) this approach can be paraphrased as follows: Differentiation of the constraint $0 = \phi(x)$ in (6) with respect to time yields an equation for λ in terms of the variable x; i.e.

$$\lambda(x) \;=\; -\{\Phi(x) J \Phi(x)^t\}^{-1}\, \Phi(x) J \nabla H(x)$$

Consider now the unconstrained Hamiltonian system

$$x' \;=\; J \nabla H_o(x) \tag{8}$$

with the modified Hamiltonian H_o

$$H_o(x) \;=\; H(x) + \phi(x)^t \lambda(x) \tag{9}$$

It is easy to see that the constraint manifold $\mathcal{M}$ is an invariant manifold of the differential equation (8) and that on $\mathcal{M}$ both (6) and (8) define the same Hamiltonian vector field. In [16], it was now suggested to integrate the unconstrained Hamiltonian system (8) instead of the formulation (6). The formulation (8) has the advantage that any symplectic integrator suitable for systems with non–separable Hamiltonian can be applied. A disadvantage of (8) is that it requires the explicit computation of λ and its derivative with respect to x. In Section 2 we will show that this reformulation of (6) as an unconstrained system is not necessary to obtain a symplectic integrator for constrained Hamiltonian systems and that (6) or (2) can be directly integrated by a symplectic IRK method. The direct integration of (6) (or (2)) has the advantage that it does not require the explicit computation of λ as a function of x. However, discretization of (8) and (6) (or (2)) by the same symplectic IRK method results in a lower order scheme for (6). This is due to the well–know order reduction for IRK methods applied to higher index DAEs [15].

Another disadvantage of the formulation (8) is that a symplectic integrator will, in general, not preserve the constraint manifold $\mathcal{M}$. The same is true for the direct integration of the formulation (6) by a symplectic IRK method. In Section 7 we will show that this is in contrast to a particular class of symplectic partitioned Runge–Kutta (PRK) methods which preserve both the symplectic structure and the constraint manifold $\mathcal{M}$.

Assume now that the discretization of (6) (or (2)) by a Runge–Kutta method results in the one–step method

$$x_{n+1} = \Psi_h(x_n) \tag{10}$$

Then Ψ_h is called symplectic if it preserves the differential 2–form given by the wedge product [2]

$$\omega^2 = dq \wedge dp = \sum_i dq^i \wedge dp^i$$

i.e. $dq_{n+1} \wedge dp_{n+1} = dq_n \wedge dp_n$, which is equivalent to

$$(\frac{d}{dx}\Psi_h(x))^t J (\frac{d}{dx}\Psi_h(x)) = J$$

We say that the method preserves the constraint $\phi(x) = 0$ if $\phi(x_n) = 0$ implies $\phi(x_{n+1}) = 0$.

Finally we mention the following useful results concerning wedge products. Let dv and du be arbitrary vectors in $\mathbf{R}^k$, $k > 0$. Then we define $dv \wedge du$ by

$$dv \wedge du \;=\; \sum_i dv^i \wedge du^i$$

where $dv^i \wedge du^i$ is the exterior product of the ith component of dv and du [2]. Let $dx = (dq, dp)$, then $dx \wedge Jdx = 2\, dq \wedge dp$. Furthermore $du \wedge Adu = 0$ for any real symmetric matrix A and $du \wedge Bdv = (B^t du) \wedge dv$ for any matrix $B \in \mathbf{R}^{k\times l}$ and $du \in \mathbf{R}^k$, $dv \in \mathbf{R}^l$. Also note that $JJ = -I$ and $J^t = -J$.

3. Symplectic Runge-Kutta Methods. The discretization of the DAE (6) by an s–stage IRK method with Butcher's tableau

$$\begin{array}{cc} c & A \\ & b^t \end{array} \tag{11}$$

leads to the system [15]

$$\begin{aligned} x_{n+1} &= x_n + h \textstyle\sum_{i=1}^s b_i J\{\nabla H(X_i) + \Phi(X_i)^t \Lambda_i\} \\ X_i &= x_n + h \textstyle\sum_{j=1}^s a_{ij} J\{\nabla H(X_j) + \Phi(X_j)^t \Lambda_j\} \qquad (i = 1, \ldots, s) \\ 0 &= \phi(X_i) \end{aligned} \tag{12}$$

(A similar system results for a discretization of (2) where $\phi(x) = g(q)$ and $\Phi(x) = G(q)$.) With (11) we associate the $s \times s$ matrix M with entries

$$m_{ij} = b_i a_{ij} + b_j a_{ji} - b_i b_j$$

The matrix M is well known from the definition of algebraic stability of RK methods [15]. Our main result is as follows.

THEOREM 3.1. *If $M = 0$ holds and A is invertible, then the method defined by (12) is symplectic.*

Proof: We use the notation

$$X_i' = J\{\nabla H(X_i) + \Phi(X_i)^t \Lambda_i\}$$

Then we have

$$dX_i \wedge JdX_i' = 0 \tag{13}$$

for all $i = 1, \ldots, s$. To see this note that

$$\begin{aligned} -dX_i \wedge JdX_i' \;=\; & dX_i \wedge \{H_{xx}(X_i)dX_i + \Phi(X_i)^t d\Lambda_i + \\ & + \sum_{k=1}^{2m} \Lambda_i^k \Gamma^k(X_i) dX_i\} \end{aligned}$$

where the components of Λ_i have been indexed by a superscript and Γ^k is the Hessian of the kth constraint function. Now H_{xx} and Γ^k are symmetric matrices and thus

$$-dX_i \wedge JdX_i' = dX_i \wedge \Phi(X_i)^t d\Lambda_i = (\Phi(X_i)dX_i) \wedge d\Lambda = 0$$

because $\phi(X_i) = 0$ implies $\Phi(X_i)dX_i = 0$.

Now differentiate the first equation in (12) and take the wedge product to arrive at

$$\begin{aligned} dx_{n+1} \wedge Jdx_{n+1} &= (dx_n + h\sum b_i dX_i') \wedge J(dx_n + h\sum b_j dX_j') \\ &= dx_n \wedge Jdx_n + 2h\sum b_i dx_n \wedge JdX_i' + \\ &\quad + h^2 \sum\sum b_i b_j dX_j' \wedge JdX_i' \end{aligned} \tag{14}$$

Now differentiate the second equation in (12) and take the wedge product of the result with JdX_i' to obtain a relation for $dx_n \wedge JdX_i'$; i.e.

$$dx_n \wedge JdX_i' = dX_i \wedge JdX_i' - h\sum a_{ij} dX_j' \wedge JdX_i'$$

Next, use this formula to eliminate $dx_n \wedge JdX_i'$ in (14). Thus getting

$$\begin{aligned} dx_{n+1} \wedge Jdx_{n+1} &= dx_n \wedge Jdx_n + 2h\sum b_i dX_i \wedge JdX_i' + \\ &+ h^2 \sum\sum (b_i b_j - 2b_i a_{ij}) dX_j' \wedge JdX_i' \end{aligned}$$

We now use (13) and the fact that

$$dX_i' \wedge JdX_j' = dX_j' \wedge JdX_i'$$

together with the assumption $m_{ij} = 0$ to derive

$$dx_{n+1} \wedge Jdx_{n+1} = dx_n \wedge Jdx_n$$

This implies the conservation of the two-form $\omega^2 = (dx \wedge Jdx)/2$. □

COROLLARY 3.2. *For constant step-size h, the implicit Gauss–Legendre RK methods result in a symplectic discretization of* (6) *and* (2). *Let s be the number of stages. Then, for* (6), *the global error satisfies*

$$x(t_n) - x_n = O(h^\nu)$$

where $\nu = s$ for s even and $\nu = s + 1$ otherwise. For (2), *we obtain*

$$q(t_n) - q_n = O(h^\nu), \qquad p(t_n) - p_n = O(h^{\nu-2})$$

where ν is as before.

Proof: For Gauss–Legendre RK methods we have $M = 0$. Order results for IRK methods applied to higher index DAEs can be found in [12]. □

REMARK 3.1. *The order of a Gauss–Legendre method with s stages applied to the unconstrained formulation* (8) *is 2s. Thus, the reformulation of a constrained Hamiltonian system as an unconstrained system has the advantage that comparable Gauss–Legendre methods lead to higher order schemes. Furthermore, higher index problems are essentially ill-posed* [18] *and special care is needed for the solution of the resulting systems of nonlinear equations. This is especially true for the index three formulation* (2).

Note, however, that the unconstrained formulation requires the explicit computation of λ and its derivative as a function of x.

As shown in [27] for unconstrained systems, higher order methods can be constructed based on the implicit midpoint rule. This idea generalizes in a straightforward way to constrained Hamiltonian systems: Let Ψ_h denote the time–h–map defined by (12) for the implicit midpoint rule. Then the composed mapping

$$\Psi_{c_1 h} \circ \Psi_{c_2 h} \circ \Psi_{c_1 h} \tag{15}$$

with $2c_1 + c_2 = 1$ and $2c_1^3 + c_2^3 = 0$ is of fourth order and symplectic. More generally, if Ψ_h is of order $p = 2k$, $k > 1$, then the composed method (15) with $2c_1 + c_2 = 1$ and $2c_1^{p+1} + c_2^{p+1} = 0$ is of order $p+2$. Methods based on the composition of second–order schemes have the computational advantage that the dimension of the systems of nonlinear equations to be solved during the integration does not increase with the order of the scheme. (Although the number of such systems of course does.)

4. Preservation of Constraints for IRK Methods. Although the method described in the previous section is symplectic for Gauss–Legendre IRK methods, it will not, in general, preserve the constraint manifold $\mathcal{M}$; i.e. $\phi(x_n) = 0$ does not imply $\phi(x_{n+1}) = 0$. To overcome this problem and the order reduction of IRK applied to index two problems, a projected IRK method was introduced in [4]. To define this method, let $\tilde{x}_{n+1}$ be given by (12), where we put $\tilde{x}_n = x_n$ and define the new x_{n+1} by

$$\begin{aligned} x_{n+1} &= \tilde{x}_{n+1} + J\Phi(x_{n+1})^t \mu_{n+1} \\ 0 &= \phi(x_{n+1}) \end{aligned}$$

However, although the projected method (12) preserves now the constraint manifold $\mathcal{M}$, the method is no longer symplectic. To see this, note that

$$dx_{n+1} \wedge J dx_{n+1} = d\tilde{x}_{n+1} \wedge J d\tilde{x}_{n+1} - d(\Phi(x_{n+1})^t \mu_{n+1}) \wedge J d(\Phi(x_{n+1})^t \mu_{n+1})$$

Thus, for the methods discussed in the previous section, we can either preserve the constraints or the symplectic structure but, in general, not both. This is even true for the particular situation of quadratic constraints $g(q) = 0$. Specifically:

It was shown by Cooper in [7] that IRK methods with $M = 0$ preserve quadratic first integrals of ordinary differential equations. In [4], a somewhat similar result was proven for the following situation: Let $g(q) = 0$ be a quadratic constraint; i.e.

$$g(q) = \frac{1}{2} q^t T q - c \tag{16}$$

where T is an arbitrary symmetric $n \times n$ matrix and c is an arbitrary positive constant. Furthermore, let us replace the index three DAE (2) by the index two formulation

$$\begin{aligned} q' &= +\nabla_p H(q,p) \\ p' &= -\nabla_q H(q,p) - G(q)^t \lambda \\ 0 &= G(q) \nabla_p H(q,p) = q^t T \nabla_p H(q,p) \end{aligned} \tag{17}$$

which is obtained from (2) by differentiating the constraint $g(q) = 0$ once. Note that g is now a first integral of this formulation. It was shown in [4] that discretization of this DAE by an IRK method with $M = 0$ results in a scheme which preserves the quadratic constraint (16); i.e., $g(q_n) = 0$ implies $g(q_{n+1}) = 0$. The crucial step in proving this is to make use of the fact that (see also [7])

$$Q_i^t T Q_i' = 0$$

for all stages $i = 1, \ldots, s$. Here Q_i' denotes

$$Q_i' = \nabla_p H(Q_i, P_i)$$

While this is true for the discretization of (17), it is easy to see that this condition is not satisfied for (12) where Q_i' is now given by

$$Q_i' = \nabla_p H(Q_i, P_i) + \phi_p(Q_i, P_i)^t \Lambda_i$$

Note furthermore that the discretization of (17) by an IRK method with $M = 0$, although it preserves quadratic constraints, does not preserve the symplectic structure (for the stage variables we get $g(Q_i) \neq 0$ and thus $dQ_i \wedge d(G(Q_i)^t \Lambda_i) \neq 0$ in general).

In Section 7, we will show that the class of partitioned Runge–Kutta methods includes methods that preserve both the sympletic structure and the constraint manifold $\mathcal{M}$.

5. Reversible Integration. The involution $R \in L(\mathbf{R}^{2n})$

$$\begin{pmatrix} \bar{q} \\ \bar{p} \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix}, \quad R = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \tag{18}$$

reverses the direction of the flow of a constrained Hamiltonian system (2) with H satisfying

$$\nabla H(x) = -R^{-1} \nabla H(Rx) \tag{19}$$

This means that time reversal $t \to -t$ together with the coordinate transformation R leave the equations (2) invariant. The standard example of a Hamiltonian satisfying (19) is provided by the separable Hamiltonian (5).

In terms of the flow ψ of the Hamiltonian system, (19) implies that ψ is R–reversible [24]; i.e.

$$\psi_{-t} = R^{-1} \psi_t R$$

or in other words that the evolution of the system backward in time is equivalent to the evolution forward in time in the changed coordinates determined by R. It is natural to ask for the same property of the time–h–map (10) of an one–step method [24]. Thus we call an IRK method R–reversible if the resulting time–h–map Ψ_h (10) satisfies

$$(\Psi_h)^{-1} = R^{-1} \Psi_h R$$

It has been shown, e.g., that the implicit midpoint rule is R–reversible for unconstrained Hamiltonian systems with the Hamiltonian H satisfying (19). The following two corollaries show that this result can be generalized to constrained Hamiltonian systems.

Corollary 5.1. *Let* (2) *be a constrained Hamiltonian system with the Hamiltonian* H *satisfying* (19). *Assume that* (2) *is discretized by a Gauss–Legendre RK method* (11) *Then the resulting scheme is* R*-reversible and symplectic.*
Proof: Gauss–Legendre RK methods (12) are symmetric and the corresponding time–h–map Ψ_h (10) satisfies

$$(\Psi_h)^{-1} = \Psi_{-h}$$

For each stage variable X_i we have

$$R^{-1}\, X_i'(R\,X_i, \Lambda_i) = -X_i'(X_i, \Lambda_i)$$

were X_i' is given by

$$X_i'(X_i, \Lambda_i) \quad = \quad J \begin{pmatrix} \nabla_p H(Q_i, P_i) \\ -\nabla_q H(Q_i, P_i) - G(Q_i)^t \Lambda_i \end{pmatrix}$$

Thus, if we apply the involution R to all the variables in (12) and premultiply all the equations by R^{-1}, then this is equivalent to replacing the stepsize h by $-h$. However, this corresponds to the inverse of Ψ_h for symmetric schemes. Therefore Ψ_h is R–reversible. The symplecticity follows from Theorem 3.1. □

The same result does not hold for the discretization of the formulation (6). This is due to the fact that the corresponding mapping to X_i' in the above proof is not R-reversible. (The exact solutions satisfy $\lambda^2 = 0$ in (7). This is no longer true for the stage variables Λ_i^2.) However, the following projection to an R-reversible mapping can be applied [24]:

Corollary 5.2. *Let* (6) *be a constrained Hamiltonian system with the Hamiltonian* H *satisfying* (19). *Assume that* (6) *is discretized by a Gauss–Legendre RK method* (11). *Denote the resulting time-*h*-map by* Ψ_h. *Then the modified time-*h*-map*

$$\tilde{\Psi}_h \quad = \quad R\,\Psi_{-h/2}\,R\Psi_{h/2}$$

is symplectic, R*-reversible, and of the same order as* Ψ_h.
Proof: $\tilde{\Psi}_h$ is symplectic if Ψ_h is. Since the flow of (6) is R-reversible, $R^{-1}(\Psi_h)^{-1}R = R\,\Psi_{-h}R$ is an approximation of the flow of the same order as Ψ_h. Finally, we see that $(\tilde{\Psi}_h)^{-1} = R^{-1}\,\tilde{\Psi}_h\,R$ because of $(\Psi_h)^{-1} = \Psi_{-h}$. □

6. Partitioned Runge–Kutta Methods. The fact that the system (1) possesses a natural partitioning suggests the use of partitioned Runge–Kutta (PRK) methods (see, e.g., [23]). In this section we show how to generalize those methods to constrained Hamiltonian systems. Our main interest in the following section is then to show that there exist PRK methods that are symplectic and preserve the constraint manifold $\mathcal{M}$.

A PRK method is specified by two tableaux

$$\begin{array}{c|c} \tilde{c} & \tilde{A} \\ \hline & \tilde{b}^t \end{array} \qquad\qquad \begin{array}{c|c} \hat{c} & \hat{A} \\ \hline & \hat{b}^t \end{array} \tag{20}$$

Straightforward application of a PRK method to the system (6) (or (2)) results in the equations

$$
(21)\quad
\begin{array}{rcl}
q_{n+1} &=& q_n + h\sum_{i=1}^{s}\tilde{b}_i\{\nabla_p H(Q_i,P_i)+\phi_p(Q_i,P_i)^t\Lambda_i\}\\
p_{n+1} &=& p_n - h\sum_{i=1}^{s}\hat{b}_i\{\nabla_q H(Q_i,P_i)+\phi_q(Q_i,P_i)^t\Lambda_i\}\\
Q_i &=& q_n + h\sum_{j=1}^{s}\tilde{a}_{ij}\{\nabla_p H(Q_j,P_j)+\phi_p(Q_j,P_j)^t\Lambda_j\}\\
P_i &=& p_n - h\sum_{j=1}^{s}\hat{a}_{ij}\{\nabla_q H(Q_j,P_j)+\phi_q(Q_j,P_j)^t\Lambda_j\} \qquad (i=1,\dots,s)\\
0 &=& \phi(Q_i,P_i)
\end{array}
$$

If we assume for now that the tableaux (20) are such that (21) indeed defines a mapping $(q_n,p_n)\to(q_{n+1},p_{n+1})$ for h small enough, then the following result can be given:

THEOREM 6.1. *Assume that the coefficients of the method* (21) *satisfy the relations*

$$\tilde{b}_i\hat{a}_{ij}+\hat{b}_j\tilde{a}_{ji}-\hat{b}_i\tilde{b}_j=0 \qquad 1\le i,j\le s \tag{22}$$

$$\tilde{b}_i=\hat{b}_i \qquad 1\le i\le s \tag{23}$$

then the method defined by (21) *is symplectic. If* (21) *is the discretization of the formulation* (2) *and if the corresponding Hamiltonian is separable; i.e. of type* (5), *then the condition* (22) *alone implies symplecticity of the method.*

Proof: The proof is analogous to that of Theorem 3.1. Note that, as in the proof of Theorem 3.1, we have $dX_i\wedge JdX_i'=0$ with $dX_i=(dQ_i,dP_i)$ and $dX_i'=(dQ_i',dP_i')$. If (21) is the discretization of (2) and the Hamiltonian is of type (5), then we even obtain $dQ_i\wedge dP_i'=dP_i\wedge dQ_i'=0$ and we can drop the condition (23). □

The same conditions on the coefficients in (20) give the symplecticity for unconstrained systems. However, (21) is the discretization of an index two DAE (index three DAE, for (2)) and thus order results for ODEs do not automatically apply to (21) [13].

7. Preservation of Constraints for PRK Methods. To begin with let us consider the discretization of (2) by the following one–stage method

$$
(24)\quad
\begin{array}{rcl}
q_{n+1} &=& q_n + h\,\nabla_p H(q_n,\bar{p}_{n+1})\\
\bar{p}_{n+1} &=& p_n - h\{\nabla_q H(q_n,\bar{p}_{n+1})+G(q_n)^t\lambda_n\}\\
0 &=& g(q_{n+1})\\[1ex]
p_{n+1} &=& \bar{p}_{n+1} - hG(q_{n+1})^t\mu_{n+1}\\
0 &=& G(q_{n+1})\nabla_p H(q_{n+1},p_{n+1})
\end{array}
$$

Let us assume that $(q_n,p_n)\in\mathcal{M}$, then Taylor expansion of $g(q_{n+1})=0$ yields

$$0=\frac{1}{2}g_{qq}(\nabla_p H,\nabla_p H)-GH_{pp}\{\nabla_q H+G^t\lambda_n\}+O(h)$$

(Here and in the sequel we frequently supress arguments when these are readily apparent.) Thus, if (3) is invertible, then (24) has a unique solution λ_n. A similar argument shows that (24) also has a unique solution μ_{n+1}. Thus (24) defines a mapping $(q_n,p_n)\in\mathcal{M}\to(q_{n+1},p_{n+1})\in\mathcal{M}$ for h small enough.

THEOREM 7.1. *The method* (24) *is symplectic, preserves the constraint manifold* $\mathcal{M}$, *and is global of first order on* $\mathcal{M}$.

Proof: The preservation of the constraint manifold $\mathcal{M}$ is obvious. The symplecticity of the method follows from the fact that

$$\begin{aligned} dq_{n+1} \wedge d\bar{p}_{n+1} &= d\,[q_n + h\,\nabla_p H(q_n, \bar{p}_{n+1})] \wedge d\bar{p}_{n+1} \\ &= dq_n \wedge d\bar{p}_{n+1} + hH_{pq} dq_n \wedge d\bar{p}_{n+1} \end{aligned}$$

$$\begin{aligned} dq_n \wedge d\bar{p}_{n+1} &= dq_n \wedge d\,[p_n - h\{\nabla_q H(q_n, \bar{p}_{n+1}) + G(q_n)^t \lambda_n\}] \\ &= dq_n \wedge dp_n - hH_{pq} dq_n \wedge d\bar{p}_{n+1} \end{aligned}$$

and

$$dq_{n+1} \wedge d(G(q_{n+1})^t \mu_{n+1}) = 0$$

Next we show that the local truncation error of (24) is of order two. To see this differentiate the constraint $g(q) = 0$ in (2) twice to obtain λ in (2) as a function of (q, p):

$$\lambda \;=\; \{GH_{pp}G^t\}^{-1}\{g_{qq}(\nabla_p H, \nabla_p H) - GH_{pp}\nabla_q H + GH_{pq}\nabla_p H\}$$

which implies that

$$\lambda_n = \lambda - \{GH_{pp}G^t\}^{-1}\{\frac{1}{2}g_{qq}(\nabla_p H, \nabla_p H) + GH_{pq}\nabla_p H\} \quad + O(h)$$

Thus the local truncation error of (24) is of order two in the variable q and of order one in the variable $\bar{p}$. For the variable μ_{n+1} we obtain the estimate

$$\mu_{n+1} \;=\; \{GH_{pp}G^t\}^{-1}\{\frac{1}{2}g_{qq}(\nabla_p H, \nabla_p H) + GH_{pq}\nabla_p H\} \quad + O(h)$$

and thus $\lambda_n + \mu_{n+1} = \lambda + O(h)$. This implies that (24) is of second order in the variable p and thus the overall method has local truncation error of order two as well. Since we assumed the Hamiltonian to be sufficiently smooth, (24) defines a smooth mapping on $\mathcal{M}$. This and standard convergence results [14] imply now that (24) is convergent of order one on $\mathcal{M}$. □

A second order scheme can be obtained from (24) by using its adjoint method [14] which is given by

$$\begin{aligned} q_{n+1} &= q_n + h\nabla_p H(q_{n+1}, \bar{p}_n) \\ \bar{p}_n &= p_n - hG(q_n)^t \mu_n \\ 0 &= g(q_{n+1}) \end{aligned}$$

$$\begin{aligned} p_{n+1} &= \bar{p}_{n+1} - h\{\nabla_q H(q_{n+1}, \bar{p}_n) + G(q_{n+1})^t \lambda_{n+1}\} \\ 0 &= G(q_{n+1})\nabla_p H(q_{n+1}, p_{n+1}) \end{aligned}$$

Denoting the time–h–map of (24) by Ψ_h and the time–h–map of the adjoint method by Ψ_h^a, we know that the composed method

$$\Psi_h^2 \;=\; \Psi_{h/2}^a \circ \Psi_{h/2} \tag{25}$$

is symmetric and is therefore of second order [14],[22].

As shown by Yoshida in [27] for unconstrained systems, one can construct now methods of arbitrarily high order. As already mentioned in Section 3, these results naturally extend to constrained systems. Specifically:

Let Ψ_h^2 denote the time–h–map of the second order method (25). Then the composed mapping

$$\Psi^2_{c_1 h} \circ \Psi^2_{c_2 h} \circ \Psi^2_{c_1 h} \tag{26}$$

with $2c_1 + c_2 = 1$ and $2c_1^3 + c_2^3 = 0$ is of fourth order, symplectic, and preserves the constraint manifold $\mathcal{M}$. More generally, if Ψ^2 is of order $p = 2k$, $k > 1$, then the composed method (26) with $2c_1 + c_2 = 1$ and $2c_1^{p+1} + c_2^{p+1} = 0$ has order $p+2$.

Note that methods constructed this way correspond to schemes (21) with $(3^k + 1)$ stages for a method of order $2k+2$. Thus, e.g., a method of order six already requires 10 stages. A 6th order method with 8 stages and a 8th order method with 16 stages were derived in [27]. These methods are also based on second order schemes and can be generalized to constrained systems.

8. Numerical Example. For the mathematical pendulum the equations (6) take the form

$$\begin{aligned}
q_1' &= p_1 + q_1\lambda_2 \\
q_2' &= p_2 + q_2\lambda_2 \\
p_1' &= \qquad -q_1\lambda_1 - p_1\lambda_2 \\
p_2' &= -g - q_2\lambda_1 - p_2\lambda_2 \\
0 &= (q_1^2 + q_2^2 - l^2)/2 \\
0 &= q_1 p_1 + q_2 p_2
\end{aligned}$$

Discretization of these equations by the implicit midpoint rule (stepsize $h = .1$ and initial value $(q_1, q_2, p_1, p_2) = (0, 1, 1, 0)$, $g = l = 1$) results in a second order symplectic method which preserves neither the constraints nor the Hamiltonian of the problem (Fig. 1–3). If the same problem is integrated by the R–reversible modification of the implicit midpoint rule (see Theorem 6.2 with $h/2$ replaced by h to keep the results comparable), then the high–frequency component in the drift gets eliminated (Fig. 3–6).

If we put $\lambda_2 = 0$ in the above DAE, then the resulting system of equations can be discretized by the PRK methods described in Section 7. In Fig. 7, the drift in the Hamiltonian H is plotted for the second order scheme (25). Note that this method preserves the constraint manifold $\mathcal{M}$ exactly.

9. Constrained Hamiltonian Systems and Lie Groups. From now on we consider constrained Hamiltonian systems with separable Hamiltonian $T(p) + V(q)$. Although the discretization of those systems can easily be discussed in terms of Runge–Kutta methods, we will follow here the Lie algebraic approach as suggested in [8], [19], and [27] for unconstrained systems with separable Hamiltonian. A main advantage of this approach is that it provides a straightforward backward error analysis [28].

In this section we generalize the necessary Lie algebraic notation to constrained Hamiltonian systems. To start with, let us introduce the Hamiltonian $H : \mathbf{R}^{2n} \times \mathbf{R}^m \to$

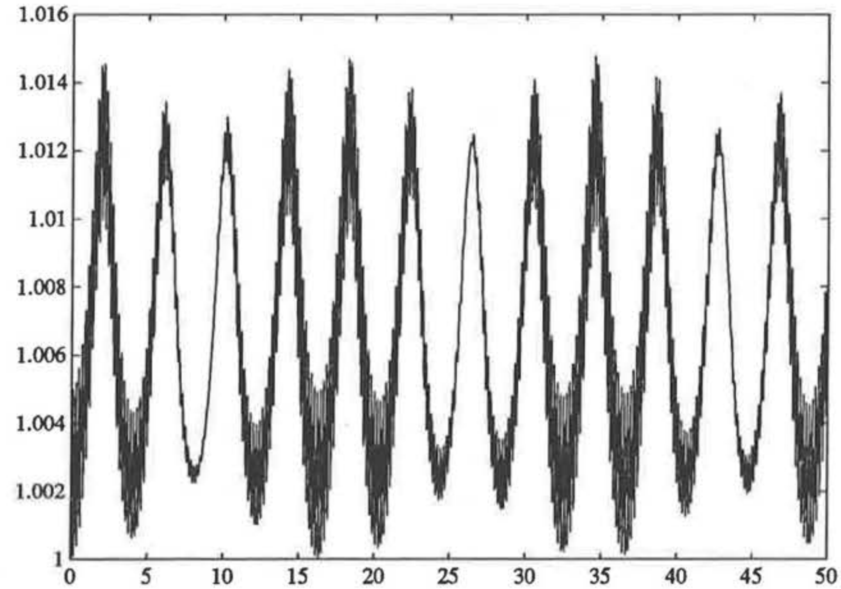

FIG. 1. *Drift in the coordinate constraint for the DAE formulation discretized by the implicit midpoint rule.*

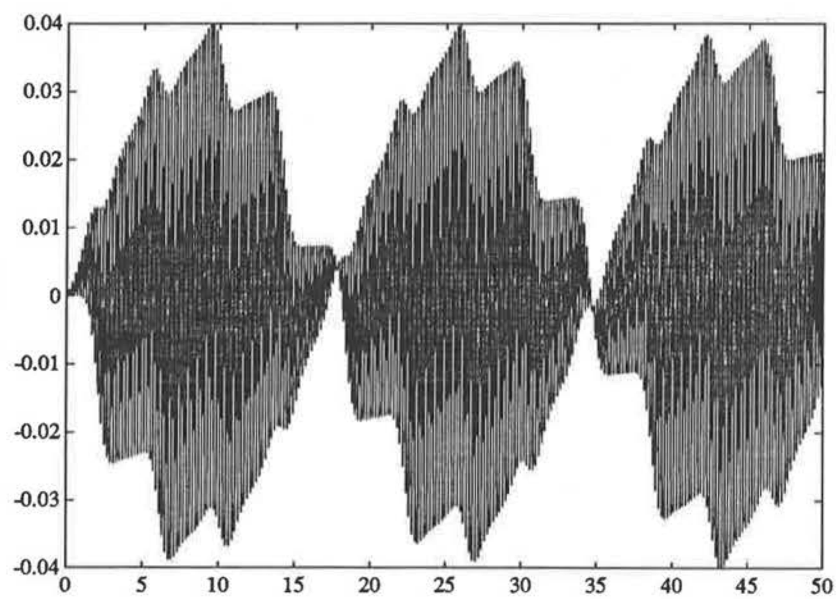

FIG. 2. *Drift in the velocity constraint for the DAE formulation discretized by the implicit midpoint rule.*

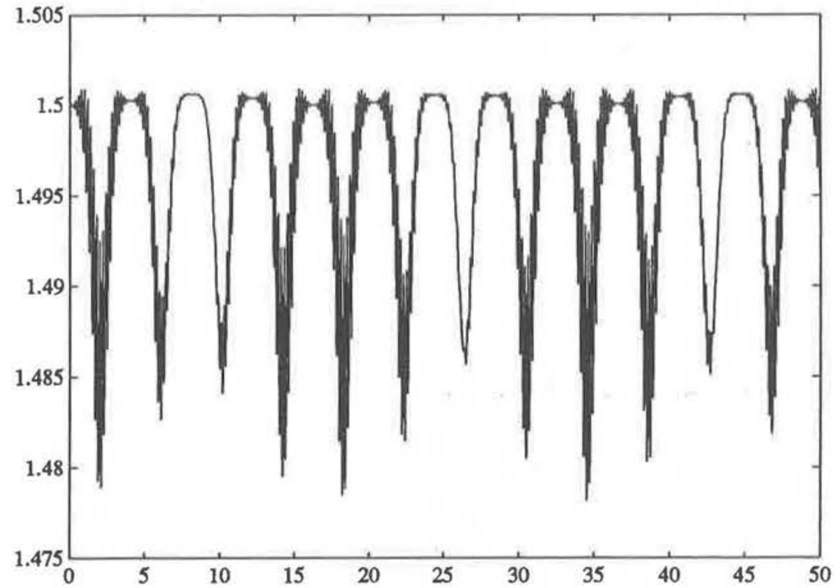

FIG. 3. *Drift in the Hamiltonian for the DAE formulation discretized by the implicit midpoint rule.*

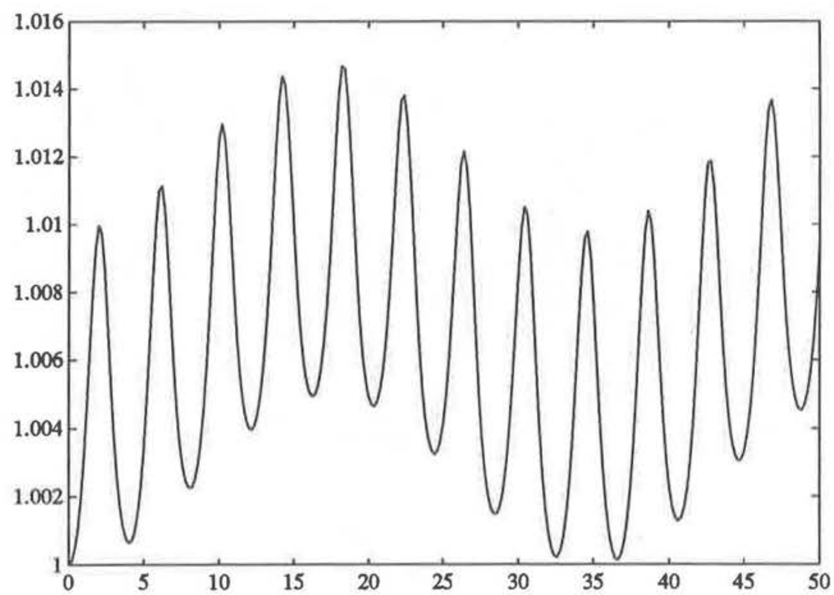

FIG. 4. *Drift in the coordinate constraint for the DAE formulation discretized by the R–reversible modification of the implicit midpoint rule.*

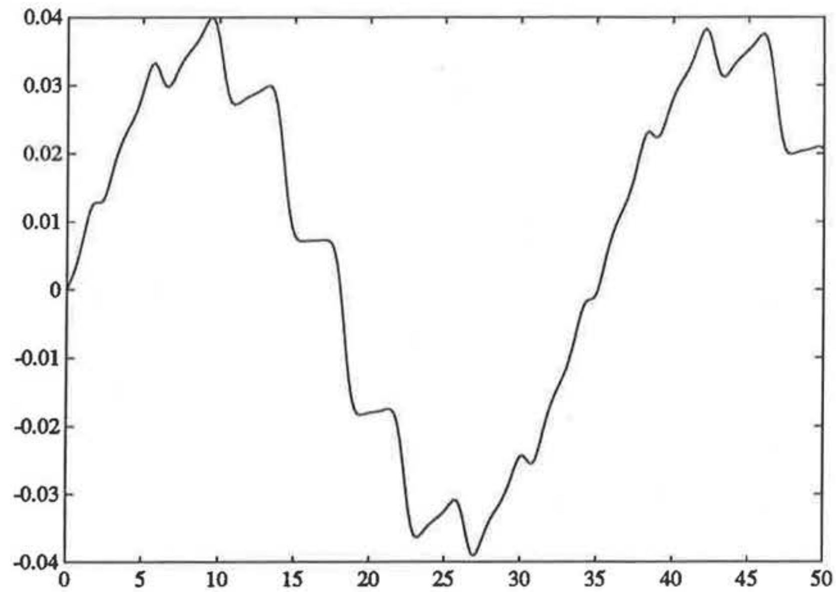

Fig. 5. *Drift in the velocity constraint for the DAE formulation discretized by the R-reversible modification of the implicit midpoint rule.*

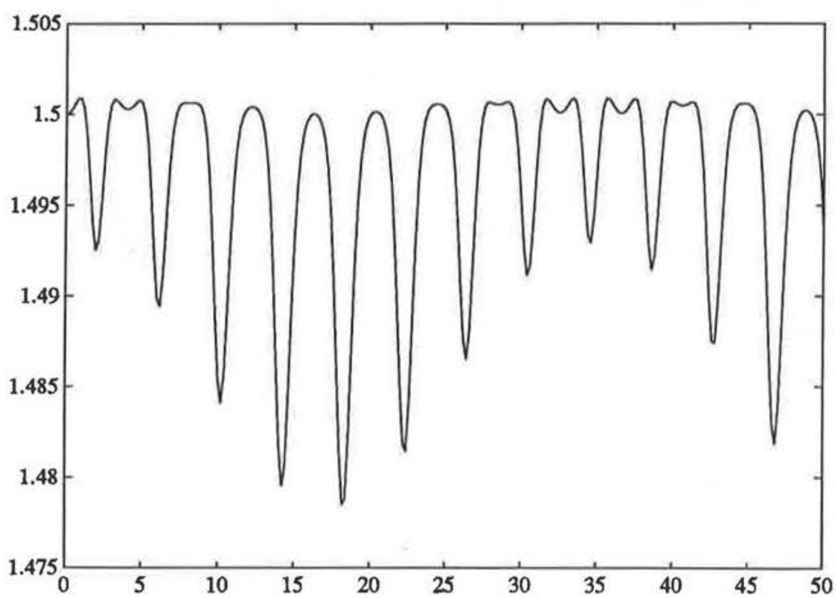

Fig. 6. *Drift in the Hamiltonian for the DAE formulation discretized by the R-reversible modification of the implicit midpoint rule.*

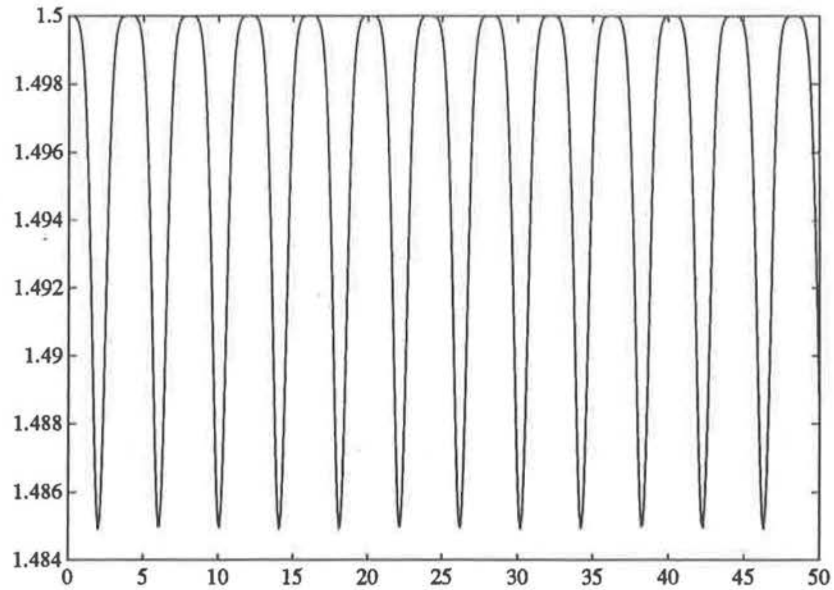

FIG. 7. *Drift in the Hamiltonian for the DAE formulation discretized by the second order PRK method.*

R

$$H(q,p\,;\lambda) \;:=\; T(p)+V(q)+g(q)^t\lambda \tag{27}$$

Then the constrained Hamiltonian system (2) can be rewritten as

$$\begin{aligned} x' &= \{x, H(\lambda)\} \\ 0 &= g(x) \end{aligned}$$

Here braces stand for the (generalized) Poisson bracket defined by

$$\{F,G\} \;=\; F_x\, J\, G_x^t$$

where J is as defined in Section 2 and $F : \mathbf{R}^{2n} \to \mathbf{R}^l$, $G : \mathbf{R}^{2n} \to \mathbf{R}^k$, $k,l \geq 1$, are sufficiently smooth functions. If we introduce the differential operator D_G, $G : \mathbf{R}^{2n} \to \mathbf{R}$, by

$$D_G\, F \;=\; \{F,G\}$$

then (2) can also be written in the form

$$\begin{aligned} x' &= D_{H(\lambda)}\, x \\ 0 &= g(x) \end{aligned} \tag{28}$$

where the Lagrange multiplier λ is implicitly determined by the constraint $g(x) = 0$ in the following way:

For $g = 0$ to be an invariant of the differential equation

$$x' \;=\; D_{H(\lambda)}\, x$$

we need that

$$0 = g' = \{g, H(\lambda)\}$$

Now

$$\{g, H(\lambda)\} \;=\; \{g, T\}$$

and thus we obtain the hidden constraint

$$f(x) := \{g, T\}(x) = G(q)\nabla T(p)$$

which again has to satisfy

$$0 = f' = \{f, H(\lambda)\}$$

This time we obtain

$$\{f, H(\lambda)\} \;=\; \{f, T\} + \{f, V\} + \{f, g\}\lambda$$

and, under the assumption that

$$\{f, g\}^{-1}(x) \;=\; (G(q)\,T_{pp}(p)\,G(q)^t)^{-1}$$

exists, this yields then

$$\lambda \;:=\; -\{f, g\}^{-1}(\{f, T\} + \{f, V\}) \tag{29}$$

Note that in the computation of λ we made use of the fact that, because $g = 0$,

$$\{F, g^t\lambda\} \;=\; \{F, g\}\lambda$$

for arbitrary functions F. Thus λ, although it is a function of x, can be formally considered as an independent parameter.

REMARK 9.1. (28) *describes a Hamiltonian vector field on the constraint manifold* $\mathcal{M}$,

$$\mathcal{M} \;=\; \{x \in \mathbf{R}^{2n} : g(x) = 0,\, f(x) = 0\}$$

As pointed out in [3], *this vector field does not depend on the values of the Hamiltonian* (27) *away from the manifold* $\mathcal{M}$. *Furthermore, the Poisson structure on* $\mathcal{M}$ *is specified by the Poisson bracket*

$$\{F, G\}_M \;=\; \{F, G\} + \{\Phi, F\}\,\{\Phi, \Phi\}^{-1}\,\{\Phi, G\}$$

where $\Phi = \phi_x$ *and the mapping* ϕ *is defined as in Section* 2*; i.e.,*

$$\phi(x) \;=\; \begin{pmatrix} g(x) \\ f(x) \end{pmatrix}$$

One can show that the restriction of the bracket $\{F, G\}_M$ *to* $\mathcal{M}$ *depends only on the restrictions of* F *and* G *to* $\mathcal{M}$. *Furthermore,* $\{g, F\}_M = \{f, F\}_M = 0$ *for arbitrary functions* $F : \mathbf{R}^{2n} \to \mathbf{R}$.

Once λ is known as a function of x, the formal solution of (28) is given by

$$x(t) \quad = \quad \exp(tD_H) \cdot x(0)$$

where $H = H(\lambda)$.

Here we used the fact that the symplectic diffeomorphisms on $\mathbf{R}^{2n}$, $\mathcal{M}$ respectively, form a Lie group. If ϕ and ψ are two elements in the group, then their product is denoted by $\phi \cdot \psi$ and the action of ϕ on an element $x \in \mathbf{R}^{2n}$ by $\phi \cdot x$. The flow generated by the differential operator D_H is denoted by $\exp(tD_H)$. The flow $\exp(tD_H)$ forms a one–parametric subgroup in the Lie group of symplectic mappings.

The time evolution of a sufficiently smooth function $F : \mathbf{R}^{2n} \to \mathbf{R}^l$ along solutions of (28) is given by

$$F' \quad = \quad D_H F \tag{30}$$

with the formal solution

$$F(x(t)) \quad = \quad \exp(tD_H) \cdot F(x(0))$$

In the following sections we will need the Taylor expansion of this formula which is given by

$$F(x(t)) \quad = \quad \sum_{i=0}^{r} \frac{t^i}{i!} D_H^i F(x(0)) + O(t^{r+1}) \tag{31}$$

where F and H are assumed to be C^r–mappings. From (30) we conclude that $F(x(t)) = F(x(0))$ if $\{F, H\} = 0$.

10. Semi–Explicit Symplectic Integrators. Let us introduce the following notation. Define $C(\lambda)$ by

$$C(\lambda) \quad = \quad g(x)^t \lambda$$

and let $D_{C(\lambda)}$ denote the corresponding differential operator. Then the exponential functions $\exp(tD_T)$, $\exp(tD_V + tD_{C(\lambda)})$, and $\exp(tD_{C(\mu)})$ can be computed explicitly by

$$\begin{aligned}
\exp(tD_T) \cdot x \quad &= \quad \begin{pmatrix} q + t\,\nabla T(p) \\ p \end{pmatrix} \\
\exp(tD_V + tD_{C(\lambda)}) \cdot x \quad &= \quad \begin{pmatrix} q \\ p - t\,(\nabla V(q) + G(q)^t \lambda) \end{pmatrix} \\
\exp(tD_{C(\mu)}) \cdot x \quad &= \quad \begin{pmatrix} q \\ p - t\,G(q)^t \mu \end{pmatrix}
\end{aligned}$$

Similar to unconstrained systems [8],[19],[27], this suggests that we should consider the mapping $\Psi_h^k : \mathcal{M} \to \mathcal{M}$, defined by

$$\Psi_h^k \quad := \quad \exp(hD_{C(\mu)}) \prod_{i=1}^{k} \exp(hc_i D_T) \cdot \exp(hd_i D_V + hd_i D_{C(\Lambda_i)}) \tag{32}$$

to approximate the exact flow

$$\exp\left(hD_{H(\lambda)}\right) : \mathcal{M} \to \mathcal{M}$$

Here the Λ_i's are implicitly determined by the fact that

$$g(x_i) \;=\; 0 \qquad (i = 1, \ldots, k) \tag{33}$$

where the x_i's are given by

$$x_i \;=\; \prod_{j=1}^{i} \exp\left(hc_j D_T\right) \cdot \exp\left(hd_j D_V + hd_j D_{C(\Lambda_j)}\right) \cdot x$$

and μ has to be chosen such that $\Psi_h^k(x) \in \mathcal{M}$.

Note that the iteration

$$x_{n+1} \;=\; \Psi_h^k \cdot x_n$$

can be implemented in the following way

$$\begin{aligned}
q_{n+1} &= Q_k \\
p_{n+1} &= P_k - hG(Q_k)^t \mu \\
0 &= G(q_{n+1})\,\nabla T(p_{n+1}) \\
\\
Q_i &= Q_{i-1} + hc_i \nabla T(P_i) \\
P_i &= P_{i-1} - hd_i(\nabla V(Q_{i-1}) + G(Q_{i-1})^t \Lambda_i) \qquad (i = 1, \ldots, k) \\
0 &= g(Q_i) \\
\\
Q_o &= q_n
\end{aligned} \tag{34}$$

which can be understood as the discretization of the DAE (2) by a k–stage partitioned Runge–Kutta method. Here $G(q) = g_q(q)$ denotes the derivative of g.

For the analysis of (32) we will use the Baker–Campbell–Hausdorff (BCH) formula [26] which states that for any two operators X and Y, the product of the exponential functions $\exp X$ and $\exp Y$ can be expressed formally as a single exponential function

$$\exp(Z) \;=\; \exp(X) \cdot \exp(Y) \tag{35}$$

where

$$\begin{aligned}
Z \;:=\; & X + Y + \frac{1}{2}[X,Y] \;+ \\
& + \frac{1}{12}([X,[X,Y]] + [Y,[Y,X]]) + \frac{1}{24}[X,[Y,[Y,X]]] \;+ \\
& - \frac{1}{720}([Y,[Y,[Y,[Y,X]]]] + [X,[X,[X,[X,Y]]]]) \;+ \\
& + \frac{1}{360}([Y,[X,[X,[X,Y]]]] + [X,[Y,[Y,[Y,X]]]]) \;+ \\
& + \frac{1}{120}([X,[X,[Y,[Y,X]]]] + [Y,[Y,[X,[X,Y]]]]) \;+ \\
& + \ldots
\end{aligned}$$

and $[X,Y] := XY - YX$, etc. Furthermore, if X and Y are differentiable operators; i.e. $X = D_A$, $Y = D_B$, then Z can be considered as a differentiable operator $Z = D_C$ with C given by

$$\begin{aligned} C &= A + B + \frac{1}{2}\{B,A\} + \\ &\quad + \frac{1}{12}(\{\{A,B\},B\} + \{\{B,A\},A\}) + \frac{1}{24}\{\{\{A,B\},B\},A\} + \\ &\quad + \dots \end{aligned}$$

Here one has to point out that, unless A and B are C^∞–mappings, only a finite number of elements in the above series exist.

We also like to mention the Jacobi identity for Poisson brackets which is given by

$$\{\{A,B\},C\} + \{\{B,C\},A\} + \{\{C,A\},B\} = 0$$

for sufficiently smooth functions $A, B, C : \mathbf{R}^{2n} \to \mathbf{R}$.

Let us start now the analysis of (32) for the simplest case $k = 1$ and $c = d = 1$. For the mapping

$$\Psi_h^1 = \exp(hD_{C(\mu)}) \cdot \exp(hD_T) \cdot \exp(hD_V + hD_{C(\Lambda)})$$

the repeated application of the BCH formula (35) yields

$$\Psi_h^1 = \exp(hD_{\bar{H}(\mu,\Lambda)}) + O(h^3)$$

where

$$\begin{aligned} \bar{H}(\mu,\Lambda) &= C(\mu) + T + V + C(\Lambda) + \\ &\quad + \frac{h}{2}(\{T, C(\mu)\} + \{V,T\} + \{C(\Lambda),T\}) \end{aligned}$$

Now, because Ψ_h^1 maps $\mathcal{M}$ into $\mathcal{M}$, we have

$$\exp(hD_{\bar{H}(\mu,\Lambda)}) \cdot g(x_n) = O(h^3)$$

and

$$\exp(hD_{\bar{H}(\mu,\Lambda)}) \cdot f(x_n) = O(h^3)$$

Since $\exp(hD_{\bar{H}})$ can be considered as the time–h–map of the flow corresponding to the Hamiltonian vector field

$$x' = \{x, \bar{H}(\mu,\Lambda)\}$$

the variables μ and Λ are determined on $\mathcal{M}$, according to (30) and (31), by

$$\{g, \bar{H}(\mu,\Lambda)\} = O(h^2)$$

and

$$\{f, \bar{H}(\mu,\Lambda)\} = O(h^2)$$

which becomes

$$0 = f - \frac{h}{2}(\{f,g\}(\Lambda - \mu) + \{f,V\}) + O(h^2)$$

and

$$0 = \{f,g\}(\mu + \Lambda) + \{f,T\} + \{f,V\} + O(h)$$

Here we used the fact that $C(\mu) = g^t\mu$, $C(\Lambda) = g^t\Lambda$, $\{g,\{V,T\}\} = -\{f,V\}$, and $\{g,T\} = f = 0$. Thus, for h small enough, μ and Λ are uniquely determined, $\mu + \Lambda = \lambda + O(h)$ and

$$\bar{H} = H + O(h)$$

on $\mathcal{M}$ where λ is as in (29). Thus we have proven the following

THEOREM 10.1. *The map Ψ_h^1 defines a first order symplectic integrator on the constraint manifold $\mathcal{M}$.*

11. Higher Order Methods. In [27], the formula

$$\exp(Z) = \exp(X) \cdot \exp(Y) \cdot \exp(X) \tag{36}$$

with

$$\begin{aligned} Z &= 2X + Y + \frac{1}{6}([Y,[Y,X]] - [X,[X,Y]]) + \\ &+ \frac{1}{360}([X,[X,[X,[X,Y]]]] - [Y,[Y,[Y,[Y,X]]]] + \\ &+ \ldots \end{aligned}$$

was used to show that symplectic integrators of arbitrarily high order exist for Hamiltonian systems of the form (1). In this section we show that this result can be generalized to constrained systems in a straightforward manner.

Let us first consider the two–stage method Ψ_h^2 as defined by (32) with $c_1 = 0$, $c_2 = 1$, $d_1 = d_2 = 1/2$, which we will rewrite as

$$\Psi_h^2 = \exp(\frac{h}{2}D_V + \frac{h}{2}D_{C(\mu)}) \cdot \exp(hD_T) \cdot \exp(\frac{h}{2}D_V + \frac{h}{2}D_{C(\Lambda)}) \tag{37}$$

Here we used the fact that the two operators D_V and D_C commute and that $D_{C(\mu)} + 1/2D_{C(\Lambda_2)}$ can formally be replaced by $1/2D_{C(\mu)}$.

COROLLARY 11.1. *The map Ψ_h^2 defines a second order symplectic integrator on the constraint manifold $\mathcal{M}$. If the Hamiltonian satisfies (19), then Ψ_h^2 is R–reversible.*

Proof: Twofold application of the BCH formula (35) to the mapping Ψ_h^2 yields

$$\Psi_h^2 = \exp(hD_{\bar{H}(\mu,\Lambda)}) + O(h^3)$$

with

$$\bar{H}(\mu,\Lambda) = T + V + \frac{1}{2}g^t(\mu + \Lambda) + \frac{h}{4}f^t(\Lambda - \mu)$$

Now using as before that on $\mathcal{M}$

$$\{g, \bar{H}(\mu,\Lambda)\} = O(h^2)$$

and

$$\{f, \bar{H}(\mu, \Lambda)\} \;=\; O(h^2)$$

we obtain $\frac{1}{2}(\mu + \Lambda) = \lambda + O(h^2)$ and $\mu - \lambda = O(h)$. Thus

$$\bar{H} \;=\; H + O(h^2)$$

on $\mathcal{M}$ which proves that (37) is of second order. The R–reversibility follows from the fact that Ψ_h^2 is symmetric. □

REMARK 11.1. *Note that the implementation of Ψ_h^2 in terms of* (21) *leads to the popular Verlet scheme with RATTLE–type constraints* [1]. *The symplecticity of this scheme has already been discussed in* [17].

Now higher–order methods can be constructed along the same lines as demonstrated in [27] for unconstrained systems. To see this note that Ψ_h^2 is a symmetric method and satisfies $\Psi_h^2\,\Psi_{-h}^2 = id$. Thus the expansion of Ψ_h^2 yields

$$\Psi_h^2 \;=\; \exp\left(h\bar{H}_1 + h^3\bar{H}_3 + h^5\bar{H}_5\right) + O(h^7)$$

This together with the formula (36) implies, for example, that the composed method

$$\Psi_{\gamma_1 h}^2 \cdot \Psi_{\gamma_2 h}^2 \cdot \Psi_{\gamma_1 h}^2$$

with $2\gamma_1 + \gamma_2 = 1$ and $2\gamma_1^3 + \gamma_2^3 = 0$ is a fourth order scheme. Since $\exp\left(hcD_V + hcD_{C(\lambda)}\right) \cdot \exp\left(hdD_V + hdD_{C(\mu)}\right)$ can be replaced formally by $\exp\left(h(c+d)D_V + h(c+d)D_{C(\lambda)}\right)$, we obtain

COROLLARY 11.2. *The mapping Ψ_h^4 in* (32) *with $c_1 = c_3 = \gamma_1$, $c_2 = \gamma_2$, $c_4 = 0$, $d_1 = d_4 = \gamma_1/2$, and $d_2 = d_3 = (\gamma_1 + \gamma_2)/2$ defines a fourth order symplectic integrator on the constraint manifold $\mathcal{M}$. Here*

$$\gamma_1 = \frac{1}{2 - 2^{1/3}}, \qquad \gamma_2 = \frac{-2^{1/3}}{2 - 2^{1/3}}$$

If the Hamiltonian satisfies (19), *then Ψ_h^4 is R–reversible.*

For the construction of 6th and 8th order methods based on the second order scheme (37) see [27].

12. Backward Error Analysis for Constrained Hamiltonian Systems. In this section we show that the solutions of the symplectic integration scheme (32) can be formally interpreted as the exact solutions of a certain perturbed constrained Hamiltonian system on the constraint manifold $\mathcal{M}$ evaluated at discrete time points. Similar results for unconstrained systems can be found, e.g., in [11], [22], [28].

Our basic assumption in this section is that the Hamiltonian (27) and the constraint function $f = \{g, T\}$ are C^r–functions, $r \geq 2$. Under this condition the λ_i's and μ in (32) are C^{r-1}–functions in $(q,p) \in \mathcal{M}$ and the Taylor expansion of Ψ_h^k with respect to h up to terms of order r is given by the BCH formula (35). Thus, as already demonstrated in the previous sections for the first and second order schemes, the time–h–map $\Psi_h^k : \mathcal{M} \to \mathcal{M}$, defined by (32), can be written as

$$\Psi_h^k \;=\; \exp\left(hD_{\bar{H}(h)}\right) + O(h^r) \tag{38}$$

where $\bar{H} : \mathcal{M} \times [0, h_o] \to \mathbf{R}$, $h_o > 0$, is a C^2–function in $(q,p) \in \mathcal{M}$.

THEOREM 12.1. *Let the Hamiltonian* (27) *and the constraint function* $f = \{g, T\}$ *be* C^r*-functions,* $r > 2$*, and let the stepsize* h *be constant,* $h < h_o$*. Then the numerical solutions* x_n*,* $n = 1, 2, \ldots, N$*, of the symplectic integration scheme* (32) *satisfy*

$$x_n = x(t_n) + O(h^r)$$

where $x(t)$ *is the exact solution of the perturbed Hamiltonian system*

$$\begin{aligned} x' &= \{x, \bar{H}(h)\}_M, \qquad x(0) = x_o \\ 0 &= g(x) \end{aligned}$$

on $\mathcal{M}$ *evaluated at the discrete time points* $t_n = nh$*. Here* $\bar{H}$ *is as in* (38) *and* $\{x, \bar{H}\}_M$ *denotes the modified Poisson bracket as defined in Remark* 9.1*. If* (32) *is a* ν*th order scheme,* $\nu < r$*, than the perturbed Hamiltonian* $\bar{H}$ *satisfies*

$$\bar{H}(h) \;=\; H + O(h^\nu)$$

on $\mathcal{M}$ *where* H *is the Hamiltonian* (27) *of the constrained system with* λ *as in* (29).
Proof: The solutions of the perturbed Hamiltonian system are given by

$$x(t) \;=\; \exp\left(t\hat{D}_{\bar{H}(h)}\right) \cdot x(0)$$

where $\hat{D}_{\bar{H}}$ is defined by $\hat{D}_{\bar{H}}\, F = \{F, \bar{H}\}_M$. As pointed out in Remark 9.1, the solutions have to satisfy $g(x(t)) = f(x(t)) = 0$.

The Λ_i's and μ in (32) are such that

$$\exp\left(hD_{\bar{H}}\right) \cdot g(x_n) \;=\; O(h^r)$$

and

$$\exp\left(hD_{\bar{H}}\right) \cdot f(x_n) \;=\; O(h^r)$$

which implies, according to (30) and (31), that $\{g, \bar{H}\} = O(h^{r-1})$ and $\{f, \bar{H}\} = O(h^{r-1})$ on $\mathcal{M}$. Thus $D_{\bar{H}} = \hat{D}_{\bar{H}} + O(h^{r-1})$ and

$$\Psi_h^k \;=\; \exp\left(h\hat{D}_{\bar{H}(h)}\right) + O(h^r)$$

Furthermore,

$$\exp\left(h\hat{D}_{\bar{H}}\right) = \exp\left(hD_H\right) + O(h^{\nu+1})$$

and therefore $\bar{H} = H + O(h^\nu)$ on $\mathcal{M}$. □

REMARK 12.1. *Following the approach taken in* [11]*, similar results could be formulated for symplectic IRK methods and general symplectic PRK methods as discussed in Sections* 3 *and* 6.

13. Acknowledgement. The current work was done while the author was visitng the University of British Columbia. The author would also like to thank Uri Ascher and Ben Leimkuhler for many fruitful discussions.

REFERENCES

[1] Anderson, H.C., RATTLE: A 'Velocity' Version of the SHAKE Algorithm for Molecular Dynamics Calculations, *J. Comp. Phys.*, 52(1983), 24–34.

[2] Arnold, V.I., *Mathematical Methods of Classical Mechanics*, Springer–Verlag, NY, 1975.

[3] Arnold, V.I., *Dynamical Systems III*, Springer–Verlag, NY, 1988.

[4] Ascher, U. and Petzold, L.R., Projected Implicit Runge–Kutta Methods for Differential–Algebraic Equations, *SIAM J. Numer. Anal.*, 28(1991)4, 1097–1120.

[5] Ascher, U. and Petzold, L.R., Stability of Computational Methods for Constrained Dynamics Systems, *SIAM J. Sci. Comput.*, 14(1993)1, 95–120.

[6] Brenan, K.E., Campbell, S.L., and Petzold, L.R., *Numerical Solution of Intial–Value Problems in DAEs*, North Holland, 1989.

[7] Cooper, G.J., Stability of Runge–Kutta Methods for Trajectory Problems, *IMA J. Numer. Anal.*, 7(1987), 1–13.

[8] Forest, E., Ruth, R.D., Fourth–Order Symplectic Integration, *Physica D*, 43(1990), 105–117.

[9] Goldstein, H., *Classical Mechanics*, Addison–Wesley, 1950.

[10] Gear, C.W., Gupta, G.K., Leimkuhler, B., Automatic Integration of Euler–Lagrange Equations with Constraints, *J. Comp. Appl. Math*, 12&13(1985), 77–90.

[11] Hairer, E., Backward Analysis of Numerical Integrators and Symplectic Methods, Technical Report, Université de Genvè, 1993.

[12] Hairer, E., Jay, L., Implicit Runge–Kutta Methods for Higher Index Differential–Algebraic Systems, Technical Report, University of Geneva, 1992.

[13] Hairer, E., Lubich, Ch., Roche, M., *The Numerical Solution of Differential–Algebraic Sustems by Runge–Kutta Methods, Lecture Notes in Mathematics*, No 1409, Springer Verlag, 1989.

[14] Hairer, E., Nørsett, S.P., Wanner, G., *Solving Ordinary Differential Equations, I*, Springer–Verlag, 1987.

[15] Hairer, E., Wanner, G., *Solving Ordinary Differential Equations, II*, Springer–Verlag, 1991.

[16] Leimkuhler, B. and Reich, S., Numerical Methods for Constrained Hamiltonian Systems, Techn. Report, Konrad–Zuse–Zentrum Berlin, 1992.

[17] Leimkuhler, B. and Skeel, R.D., Symplectic Numerical Integrators in Constrained Hamiltonian Systems, Techn. Report, Univ. of Kansas, Lawrence, 1992.

[18] März, R., Numerical Methods for Differential–Algebraic Equations, *Acta Numerica*, 1992, 141–198.

[19] Neri, F., Lie Algebras and Canonical Integration, Department of Physics, University of Maryland, preprint, 1988.

[20] Okunbor, D. and Skeel, R.D., Explicit canonical methods for Hamiltonian Systems, *Math. Comp.*, 59(1992), 439–455.

[21] Ruth, R.D., A Canonical Integration Technique, *IEEE Trans. Nucl. Sci.*, 30(1983), 2669–2671.

[22] Sanz–Serna, J.M., Symplectic Integrators for Hamiltonian Problems: An Overview, *Acta Numerica*, 1(1992), 243–286.

[23] Sanz–Serna, J.M., Runge–Kutta Schemes for Hamiltonian Systems, *BIT*, 28(1988), 877–883.

[24] Scovel, C., Symplectic Integration of Hamiltonian Systems, in *The Geometry of Hamiltonian Systems*, Ratiu, T. (editor), Springer Verlag, 1991.

[25] Tobias, D.J., Brooks, C.L., Molecular Dynamics with Internal Coordinate Constraints. *J. Chem. Phys.*, 89(1988), 5115–5127.

[26] Varadarajan, V.S., *Lie Groups, Lie Algebras, and Their Reprsentation*, Prentice–Hall, Englewood Cliffs, 1974.

[27] Yoshida, H., Construction of Higher Order Symplectic Integrators, *Phys. Lett. A*, 150(1990), 262–268.

[28] Yoshida, H., Recent Progress in the Theory and Application of Symplectic Integrators, to appear in *Celestial Mech.*, 1993.