On higher-order semi-explicit symplectic partitioned Runge-Kutta methods for constrained Hamiltonian systems

Sebastian Reich
Konrad-Zuse-Zentrum
Heilbronner Straße 10
D–10711 Berlin, Germany
e–mail: na.reich@na-net.ornl.gov

January 31, 1996

Abstract

In this paper we generalize the class of explicit partitioned Runge-Kutta (PRK) methods for separable Hamiltonian systems to systems with holonomic constraints. For a convenient analysis of such schemes, we first generalize the backward error analysis for systems in R^m to systems on manifolds embedded in R^m. By applying this analysis to constrained PRK methods, we prove that such methods will, in general, suffer from order reduction as well-known for higher-index differential-algebraic equations. However, this order reduction can be avoided by a proper modification of the standard PRK methods. This modification increases the number of projection steps onto the constraint manifold but leaves the number of force evaluations constant. We also give a numerical comparison of several second, fourth, and sixth order methods.

Keywords: Hamiltonian systems, holonomic constraints, differential-algebraic equations, symplectic discretization, partitioned Runge-Kutta methods, composition methods, backward error analysis

AMS(MOS) Subject Classifications: 65L05, 65L20

1 Introduction

In this paper we discuss the generalization of explicit symplectic methods for separable Hamiltonian systems of type

$$\begin{array}{rcl} q' &=& +\nabla T(p)\,, \\ p' &=& -\nabla V(q) \end{array} \tag{1}$$

with $p, q \in R^n$ and $V, T : R^n \to R$ to systems with holonomic constraints; i.e.,

$$\begin{array}{rcl} q' &=& +\nabla T(p)\,, \\ p' &=& -\nabla V(q) - G(q)^T\lambda\,, \\ 0 &=& g(q) \end{array} \tag{2}$$

where $g : R^n \to R^l$, $1 \leq l < n$, is the vector-valued constraint function and $G(q) \in R^{l\times n}$ its Jacobian [2]. In most applications $T(p)$ represents kinetic energy and is of the form

$$T(p) \;=\; \frac{p^T M^{-1} p}{2}\,.$$

The equations (2) constitute a differential-algebraic system (DAE) of index three with the hidden constraint

$$0 = G(q)\nabla T(p)$$

which is obtained by taking the time derivative of $g(q) = 0$. One further differentiation with respect to time yields then an implicit equation for the Lagrange multiplier λ in (2). This equation can be solved for λ if the matrix $G(q)T_{pp}(p)G(q)^T$ is invertible which we will assume from now on. Thus (2) defines a Hamiltonian system on the constraint manifold

$$\mathcal{M} := \{(q,p) : g(q) = 0,\, G(q)\nabla T(p) = 0\}.$$

The symplectic integration of (2) has been discussed before by Leimkuhler and Reich [7], Leimkuhler and Skeel [8], Reich [13], Jay [6], Reich [15], and McLachlan and Scovel [10]. It has been shown, for example, that the partitioned Runge-Kutta (PRK) methods consisting of the pair of Lobatto IIIA and Lobatto IIIB methods can be generalized to (2) such that they are symplectic, constraint preserving, and of the same order as for the system (1) [6]. However, the resulting schemes are implicit in the force field $-\nabla V$ for order $p > 2$ and require significant more storage than, for example, explicit PRK methods. In [15], it has been shown that any first or second order symplectic integrator for unconstrained systems can be generalized to (2) such that it is symplectic and constraint preserving. Based on this result, higher-order methods can be obtained by a proper composition of second order methods as suggested by Yoshida [17] for unconstrained systems [15]. In this paper, we will look at the class of explicit PRK methods and discuss its generalization to constrained systems. We will see that, although the formal generalization is straightforward, the resulting method will, in general, suffer from order reductions as well known for higher index DAEs. To prove this result, we generalize the standard backward error analysis (see, e.g., [4]) for unconstrained systems to systems with constraints. This analysis implies that composition methods based on first or higher order constrained methods will not suffer from order reduction. Thus, using a result by McLachlan on the composition of unconstrained first and second order methods [9], one can generalize any s-stage explicit PRK of order p to a semi-explicit PRK for constrained systems (2) of the same order p with, in general, $2s-1$ stages. The resulting scheme requires $2s-1$ projections onto the constraint $g(q) = 0$ but only s evaluations of $\nabla V(q)$ per integration step. One can do better for PRK methods based on the composition of symmetric PRK methods of order two. Here any s-stage PRK method translates directly into a s-stage PRK method for constrained systems without order reduction. In Section 7, we compare several second, fourth, and sixth order methods by numerical integration of the double pendulum.

2 Semi-explicit PRK methods for constrained systems

Any explicit symplectic PRK method for the system (1) can be written in the form [5], [11]

$$\begin{aligned}
Q_0 &= q_n\,,\\
P_0 &= p_n\,,\\
&\text{for } i = 1,2,\ldots,s\\
&\qquad P_i = P_{i-1} - h\tilde{b}_i\nabla V(Q_{i-1})\,,\\
&\qquad Q_i = Q_{i-1} - h\tilde{B}_i\nabla T(P_i)\,,\\
q_{n+1} &= Q_s\,,\\
p_{n+1} &= P_s
\end{aligned} \tag{3}$$

with properly chosen coefficients $(\tilde{b}_i)$ and $(\tilde{B}_i)$. As already suggested in [13], it seems natural to generalize this class of PRK methods to constrained systems (2) in the following way:

$$\begin{aligned}
Q_0 &= q_n\,,\\
P_0 &= p_n\,,\\
&\text{for } i = 1,2,\ldots,s\\
&\qquad \begin{aligned}
P_i &= P_{i-1} - hb_i\nabla V(Q_{i-1}) - h\,G(Q_{i-1})^T\lambda_i)\,,\\
Q_i &= Q_{i-1} + hB_i\nabla T(P_i)\,,\\
0 &= g(Q_i)\,,
\end{aligned}\\
q_{n+1} &= Q_s\,,\\
p_{n+1} &= P_s - h\,G(Q_s)^T\mu\,,\\
0 &= G(q_{n+1})\nabla T(p_{n+1})\,.
\end{aligned} \tag{4}$$

We also allow for $b_i = 0$ or $B_i = 0$. In case $B_i = 0$, we simply set $\lambda_i = 0$.

This scheme (4) is clearly constraint preserving. It also has been shown [13] to be symplectic. However, it is not obvious that the choice $(b_i) = (\tilde{b}_i)$ and $(B_i) = (\tilde{B}_i)$ implies that (4) has the same order as (3). In fact, according to results given by Ostermann [12] for semi-explicit methods for DAEs of index three, one has to expect order reduction, in general. This will be made explicit for (4) in Section 5.

Remark. For $T(p) = (p^T M^{-1} p)/2$, the method (4) can also be considered as the natural extension of explicit Runge-Kutta-Nyström methods [5], [11] to constrained systems. Thus there is no need to consider those methods separately.

3 Mathematical background

Upon using the Lie algebraic formalism [16], one can rewrite (3) as

$$\Psi_h = \prod_{i=1}^{s} \exp\left(h\tilde{B}_i X_T\right)\cdot\exp\left(h\tilde{b}_i X_V\right) \tag{5}$$

and

$$(q_{n+1},p_{n+1}) = \Psi_h\,(q_n,p_n)\,.$$

Here $\exp(hX_V)$ and $\exp(hX_T)$ denote the exact time-h-flows of the differential equations

$$\begin{aligned}
q' &= 0\,,\\
p' &= -\nabla V(q)\,,
\end{aligned}$$

and

$$\begin{aligned}
q' &= \nabla T(p)\,,\\
p' &= 0\,.
\end{aligned}$$

By applying the Baker-Campbell-Hausdorff (BCH) formula [16], we can rewrite (5) as

$$\Psi_h = \prod_{i=1}^{s} \exp(hX_{\hat{H}_i}(h)) \tag{6}$$

where $\hat{H}_i(h)$ is given by

$$\hat{H}_i(h) = \tilde{b}_i V + \tilde{B}_i T + \frac{h}{2}\tilde{b}_i\tilde{B}_i\{V,T\} + \frac{h^2}{12}(\tilde{b}_i^2\tilde{B}_i\{V,\{V,T\}\} + \tilde{b}_i\tilde{B}_i^2\{T,\{T,V\}\}) + O(h^3)$$

and $\{F,H\}$ denotes the Poisson bracket of two functions $F,G : R^{2n} \to R$; i.e.

$$\{F,G\} = F_q G_p^T - F_p G_q^T .$$

The vector field $X_{\hat{H}_i}$ is defined by

$$X_{\hat{H}_i}(h) := \{\mathrm{id}, \hat{H}_i(h)\}$$

where id denotes the identity map in R^{2n}; i.e. $\mathrm{id}(q,p) = (q,p)$, and $\exp(tX_{\hat{H}_i}(h))$ denotes the exact time-t-flow corresponding to $X_{\hat{H}_i}(h)$.

Further application of the BCH formula to (6) yields

$$\Psi_h = \exp(hX_{\hat{H}}(h))$$

where

$$\hat{H}(h) = \alpha_1 V + \alpha_2 T + \frac{h}{2}\alpha_3\{V,T\} + \frac{h^2}{12}(\alpha_4\{V,\{V,T\}\} + \alpha_5\{T,\{T,V\}\}) + O(h^3) \qquad (7)$$

with well defined coefficients (α_i). It is obvious from this formula that, for example, there are five order conditions for a PRK method of order three; namely $\alpha_1 = \alpha_2 = 1$ and $\alpha_3 = \alpha_4 = \alpha_5 = 0$, and four conditions for a symmetric PRK method of order four; namely $\alpha_1 = \alpha_2 = 1$ and $\alpha_4 = \alpha_5 = 0$. (Note that, for symmetric methods, all terms in $\hat{H}$ of odd order in h disappear automatically.)

Let ϕ_h^+ denote the first order symplectic integrator

$$\phi_h^+ = \exp(hX_T)\cdot\exp(hX_V)$$

and ϕ_h^- its adjoint method defined by

$$\phi_h^- = \exp(hX_V)\cdot\exp(hX_T).$$

Then one can rewrite (5) as

$$\Psi_h = \prod_{i=1}^{s}\phi_{d_i h}^- \cdot \phi_{c_i h}^+ \qquad (8)$$

with properly defined coefficients (c_i) and (d_i) [9]. Since $d_s = 0$, the resulting scheme can now be considered as an explicit PRK method with $(2s-1)$ stages. Of course, the scheme is reducible. However, as shown by McLachlan in [9], the method (8) is of the same order as (5) even if we replace ϕ_h^+ by an arbitrary first order integrator of (1) and ϕ_h^- by the corresponding adjoint method. As already formally suggested by Reich [13] and McLachlan and Scovel [10], this approach yields another possible generalization of PRK methods (3) to constrained systems (2). This will be discussed in more detail in the following section and Section 6.

4 Backward error analysis for integrators on manifolds

Before analyzing the constrained PRK method (4), we will derive a few general results on the numerical integration of C^∞-vector fields $X : \mathcal{M} \to T\mathcal{M}$,

$$x' = X(x) \tag{9}$$

on a C^∞-manifold $\mathcal{M} \subset R^m$.

We denote the flow of (9) by $\exp(tX)$; i.e.

$$x(t) = \exp(tX)\, x(0)\,.$$

Let

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

with $\Psi_h : \mathcal{M} \to \mathcal{M}$ be a numerical approximation of the exact time-h-flow of (9). We want to show that there is a perturbed vector field $\hat{X}(h) : \mathcal{M} \to T\mathcal{M}$ such that the corresponding flow $\exp(t\hat{X}(h))$ satisfies

$$\Psi_h = \exp(h\hat{X}(h))$$

up to terms of arbitrarily high order in h. Furthermore, if X is Hamiltonian and Ψ_h is a symplectic map, then $\hat{X}(h)$ is Hamiltonian as well. These properties have been proven before for numerical integrators in R^m (see, for example, [4], [14],[3]).

To generalize these results to numerical schemes on manifolds, we assume that there exists a C^∞-mapping $\rho : R^m \to R^k$ such that

$$\mathcal{M} = \{x \in R^m : \rho(x) = 0\}$$

and the Jacobian $D\rho$ of ρ has full rank on $\mathcal{M}$. As we will see below, this allows one to reformulate (9) and (10) on an open neighborhood U of $\mathcal{M}$ in R^m with $\mathcal{M}$ as an invariant manifold. Now we can first apply the standard backward error analysis for ODEs to this extended formulation and then reduce the perturbed ODE back to the manifold $\mathcal{M}$ which yields the appropriate perturbed vector field for the scheme (10) on $\mathcal{M}$ (see Figure 1).

Specifically: Let W be a sufficiently small open neighborhood of the zero element in R^k. Then we define a neighborhood $U \subset R^m$ of $\mathcal{M}$ in R^m through

$$U = \bigcup_{\epsilon \in W} \{x \in R^m : \rho(x) = \epsilon\}\,.$$

Note that U is C^∞-diffeomorph to $\mathcal{M} \times W$; i.e., $x \in U \longleftrightarrow (z, \epsilon) \in \mathcal{M} \times W$ where $x := z + D\rho(z)^T\lambda$ with $\lambda \in R^k$ defined through

$$\rho(z + D\rho(z)^T\lambda) = \epsilon\,. \tag{11}$$

Let us denote this diffeomorphism by $\phi : \mathcal{M} \times W \to U$. Now any mapping ψ on $\mathcal{M}$ can be extended to a mapping on U by first defining the mapping $\psi \times \mathrm{id}$ on $\mathcal{M} \times W$ and subsequent transformation of this mapping to a mapping on U via the above described diffeomorphism between $\mathcal{M} \times W$ and U. Let us denote this extension by $\psi^e : U \to U$; i.e.,

$$\psi^e = \phi \cdot (\psi \times \mathrm{id}) \cdot \phi^{-1}\,.$$

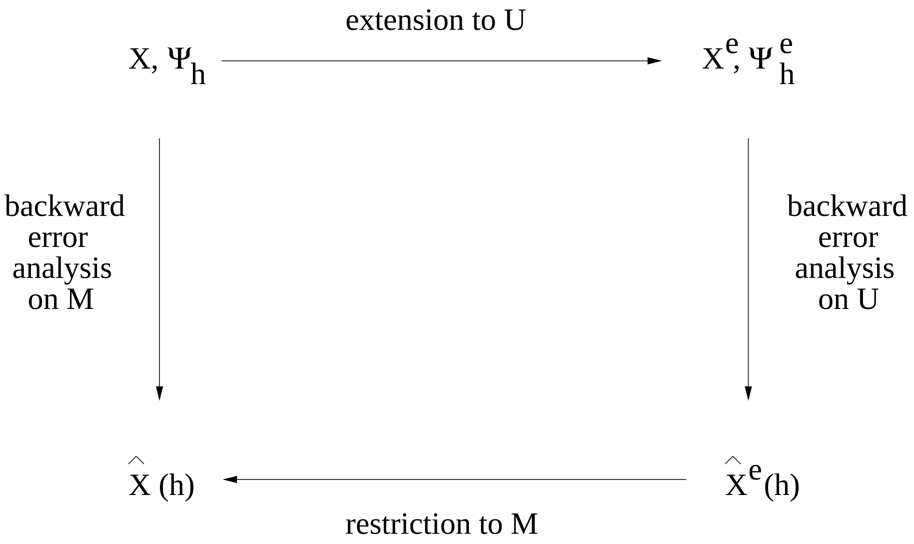

Figure 1: Commutative diagram for backward error analysis.

By definition we have $\psi^e(x) = \psi(x)$ for all $x \in \mathcal{M}$. Of course, the same idea applies to a one-parametric family ψ_h. For a vector field X on $\mathcal{M}$ we obtain the extended vector field $X^e : U \to TU$ defined by

$$X^e \;=\; T\phi \cdot (X \times T\mathrm{id}) \cdot \phi^{-1}$$

where $T\phi : T\mathcal{M} \times TW \to TU$ denotes the tangential map of ϕ [2] and $T\mathrm{id} : W \to TW$ is the zero vector field on W. Note that we have $X^e(x) = X(x)$ for all $x \in \mathcal{M}$ and that the mapping ρ is a first integral of X^e on U; i.e., $L_{X^e}\rho \equiv 0$. Here L_{X^e} denotes the Lie derivative of the vector field X^e [2].

We are now ready to state our

Theorem 1. Let Ψ_h be a numerical approximation to the exact time-h-flow $\exp(hX)$ of (9) with local discretization error of order p. Then there exists a perturbed vector field $\hat{X}(h) : \mathcal{M} \to T\mathcal{M}$ (in terms of an asymptotic series in h) such that its flow $\exp(t\hat{X}(h))$ satisfies

$$\Psi_h \;=\; \exp(h\hat{X}(h))$$

up to terms of arbitrarily high order in h and

$$\hat{X}(h) \;=\; X + O(h^p)\,.$$

Proof. We follow the commutative diagram of Figure 1; i.e., we first extend the one-parametric families Ψ_h and $\exp(hX)$ as well as the vector field X to a neighborhood U of $\mathcal{M}$ in $\boldsymbol{R}^m$. Then the standard backward error analysis is applied to Ψ_h^e which yields a perturbed vector field $\hat{X}^e(h)$ on U

with $\hat{X}^e(h) = X^e + O(h^p)$. Now, since the mapping ρ is an invariant of Ψ^e as well as a first integral of X^e, it is easy to show [14] that ρ is also a first integral of the perturbed vector field $\hat{X}^e(h)$. Thus we can restrict $X^e(h)$ to $\mathcal{M}$ and obtain a perturbed vector field $\hat{X} : \mathcal{M} \to T\mathcal{M}$. This vector field is uniquely determined by the mappings Ψ_h and $\exp(hX)$ and does not depend on the particular diffeomorphism ϕ used for the extension. Thus $\Psi_h = \exp(h\hat{X}(h))$ and $\hat{X}(h) = X + O(h^p)$ as claimed. In particular, as for differential equations in R^m, the perturbed vector field can be written as

$$\hat{X}(h) \;=\; X + \sum_{k=p}^{\nu} \frac{h^k}{(k+1)!} X^{[k+1]},$$

$\nu \gg p$, $X^{[k+1]} : \mathcal{M} \to T\mathcal{M}$. □

We can now also generalize the BCH-formula to the composition of flows on $\mathcal{M}$. Let L_X denote the Lie derivative of a vector field $X : \mathcal{M} \to T\mathcal{M}$ and let $[X,Y]_{\mathcal{M}} = L_Y L_X - L_X L_Y$ denote the Lie-Jacobi bracket of two vector fields X and Y on $\mathcal{M}$ [2]. Note that $[.,.]_{\mathcal{M}}$ is equivalent to the restriction of the standard Lie-Jacobi bracket $[.,.]$ on R^m to $\mathcal{M}$ [2]. Then we obtain

Corollary 1. Let X and Y be two vector fields on $\mathcal{M}$ with corresponding flows $\exp(tX)$ and $\exp(tY)$ respectively. Then there is a perturbed vector field $\hat{Z}(h)$ such that

$$\exp(h\hat{Z}(h)) \;=\; \exp(hY)\cdot\exp(hX) \tag{12}$$

up to terms of arbitrarily high order. The expansion of $\hat{Z}(h)$ is given by

$$\hat{Z}(h) \;=\; X + Y + \frac{h}{2}[X,Y]_{\mathcal{M}} + \frac{h^2}{12}([X,[X,Y]_{\mathcal{M}}]_{\mathcal{M}} + [Y,[Y,X]_{\mathcal{M}}]_{\mathcal{M}}) + O(h^3).$$

Proof. Again we first reformulate the problem in terms of the standard BCH-formula on U. Now, since ρ is a first integral of X^e and Y^e, the standard Lie-Jacobi bracket $[X^e, Y^e]$ on U is equivalent to $[X^e, Y^e]_{\mathcal{M}}$ where, with a slight abuse of notation, $[.,.]_{\mathcal{M}}$ stands for the extension of the Lie-Jacobi bracket $[.,.]_{\mathcal{M}}$ on $\mathcal{M}$ to U. Finally we restrict the resulting expansion to $\mathcal{M}$. □

Theorem 1 and Corollary 1 allow one to generalize the composition method (8) to constrained systems.

Corollary 2. Let $\phi_h^+ : \mathcal{M} \to \mathcal{M}$ be a first order integrator of (9) and let $\phi_h^- : \mathcal{M} \to \mathcal{M}$ denote its adjoint; i.e. $\phi_{-h}^+ = (\phi_h^-)^{-1}$. Assume that (8) with coefficients (c_i) and (d_i) is of order p for unconstrained systems. Then (8) is of the same order for the first order integrators ϕ_h^+ and ϕ_h^- on $\mathcal{M}$.

Proof. The perturbed vector fields for ϕ_h^+ and ϕ_h^- satisfy

$$\hat{X}^+(h) \;=\; X + \sum_{k=1}^{\nu} \frac{(+h)^k}{(k+1)!} X^{[k+1]}$$

and

$$\hat{X}^-(h) \;=\; X + \sum_{k=1}^{\nu} \frac{(-h)^k}{(k+1)!} X^{[k+1]},$$

$\nu > p$. This is analog to the situation for unconstrained systems. The order conditions for the coefficients (c_i) and (d_i) can now be obtained by a recursive application of (12) and we obtain the same order conditions as for unconstrained systems. □

Remark. Following the argumentation of Corollary 2, one could now also give a proof for the generalization of Yoshida's composition methods [17] to constrained systems.

So far we have considered arbitrary vector fields on $\mathcal{M}$. Let us now consider the subalgebra of Hamiltonian vector fields on $\mathcal{M}$. The corresponding group is the group of symplectic diffeomorphisms on $\mathcal{M}$ [2]. Of course, we have to assume that $\mathcal{M}$ is a symplectic manifold with Lie-Poisson bracket $\{.,.\}_{\mathcal{M}}$ which implies that a Hamiltonian function H generates the Hamiltonian vector field $X_H = \{\mathrm{id}, H\}_{\mathcal{M}}$. We also assume that $\{.,.\}_{\mathcal{M}}$ is the restriction of the canonical Lie-Poisson structure of R^m, m even, to $\mathcal{M}$. Note that the Lie-Poisson bracket $\{.,.\}_{\mathcal{M}}$ satisfies $[X_F, X_G]_{\mathcal{M}} = X_H$ with $H = \{G, F\}_{\mathcal{M}}$. Note also that the Lie-Poisson bracket $\{.,.\}_{\mathcal{M}}$ on $\mathcal{M} \in R^{2n}$ can be extended to a neighborhood U of $\mathcal{M}$ in R^m [2] through

$$\{F,G\}_{\mathcal{M}} \;\; = \;\; \{F,G\} - \{F,\rho\}\{\rho,\rho\}^{-1}\{\rho,G\}$$

where F and G are any two functions defined on U. It can be shown that the function $\{F,G\}_{\mathcal{M}}$ restricted to $\mathcal{M}$ depends only on the values of F and G on $\mathcal{M}$. For more details on systems with holonomic constraints $g(q) = 0$ see Section 5.

The proper extension of Theorem 1 to the class of Hamiltonian vector fields on $\mathcal{M}$ is given by

Theorem 2. Let (9) be a Hamiltonian vector field with Hamiltonian H. We assume that the numerical approximation Ψ_h of the exact time-h-flow is a symplectic map. Then there is a perturbed Hamiltonian function $\hat{H}(h) : \mathcal{M} \to R$ such that the corresponding Hamiltonian vector field $X_{\hat{H}}(h) = \{\mathrm{id}, \hat{H}(h)\}_{\mathcal{M}}$ satisfies the results of Theorem 1.

Proof. In contrast to Theorem 1 we have now to require that the transformation ϕ is a canonical mapping between $\mathcal{M} \times W$ and U. (The Lie-Poisson structure on U is the canonical structure $\{.,.\}$ of R^m, m even, while $\{.,.\}_{\mathcal{M}}$ is the Lie-Poisson structure on $\mathcal{M}$ and W has the canonical structure of R^k, k even). Note that the transformation ϕ defined by (11) is not canonical. However, a canonical transformation ϕ can always be found at least locally (this is insured by Darboux' Theorem [2]). With ϕ a canonical transformation, the extended one-parametric family Ψ_h^e is symplectic on U provided Ψ_h is symplectic on $\mathcal{M}$. The corresponding statement holds for the extended vector field X^e. Thus $\hat{X}^e(h)$ is Hamiltonian and so is the restricted vector field $\hat{X}(h)$. □

In the following two sections we will use these results to investigate the order conditions for (4).

5 Order reduction for constrained PRK methods

Using results of [15], one can rewrite (4) similar to (3) as

$$\Psi_h \;\; = \;\; \exp\left(hX_{g^T\mu}\right) \prod_{i=1}^{s} \exp\left(hB_iX_T\right) \cdot \exp\left(hb_iX_V\right) \cdot \exp\left(hX_{g^T\lambda_i}\right) \tag{13}$$

where the λ_i's are chosen such that $g(Q_i) = 0$ and μ such that the hidden constraint function

$$f(q,p) \;\; := \;\; G(q)\nabla T(p)$$

is satisfied at the endpoint; i.e., $f(q_{n+1}, p_{n+1}) = 0$.

It is worthwhile to note that

$$\begin{aligned}\{V, g^t\lambda_i\} &= \{V, g\}\lambda_i + \{V, \lambda_i\}g \\ &= 0\end{aligned}$$

for $g = 0$. Thus the two vector fields X_V and $X_{g^T\lambda_i}$ commute on the constraint manifold $\mathcal{M}$.

To simplify the analysis of (4), we modify (13) such that the hidden constraint $f = 0$ is also satisfied after each substep i. This can be achieved by using

$$\Psi_h = \prod_{i=1}^{s} \exp\left(hX_{g^T\mu_i}\right) \cdot \exp\left(hB_iX_T\right) \cdot \exp\left(hb_iX_V\right) \cdot \exp\left(hX_{g^T\lambda_i}\right) \tag{14}$$

where the μ_i's are chosen such that $f(Q_i, P_i) = 0$ in (4). Note that this modification does not change the overall map Ψ_h since the effect of $\exp\left(hX_{g^T\mu_{i-1}}\right)$ gets canceled by the subsequent application of $\exp\left(hX_{g^T\lambda_i}\right)$ in the next substep.

Let us write down the perturbed Hamiltonian function for each substep i in (14). By applying the standard BCH formula [16], we obtain

$$\hat{H}_i(h) = H_i^{[1]} + \frac{h}{2}H_i^{[2]} + \frac{h^2}{12}H_i^{[3]} + O(h^3)$$

with

$$H_i^{[1]} := b_iV + B_iT + g^T(\lambda_i + \mu_i)\,,$$

$$H_i^{[2]} := b_iB_i\{V, T\} + B_i\, f^T(\lambda_i - \mu_i)\,,$$

and

$$\begin{aligned}H_i^{[3]} := {} & b_iB_i^2\{T, \{T, V\}\} + b_i^2B_i\{V, \{V, T\}\} + B_i^2\{T, f\}(\lambda_i - \mu_i) + \\ & b_iB_i\{V, f\}(2\lambda_i - 4\mu_i) + B_i(\lambda_i - \mu_i)^T\{g, f\}(\lambda_i - \mu_i) - 2B_i\mu_i^T\{g, f\}\lambda_i\end{aligned}$$

where we used $f = \{g, T\}$. (Formally, $H_i^{[3]}$ consists of further terms that contain f and g outside Poisson brackets. Since those terms are equal to zero on $\mathcal{M}$, we dropped them here.)

Now, since g and f are preserved, we have

$$\{g, \hat{H}_i(h)\} = 0$$

and

$$\{f, \hat{H}_i(h)\} = 0\,.$$

Thus

$$\lambda_i + \mu_i = -\{g, f\}^{-1}(B_i\{T, f\}^T + b_i\{V, f\}^T) + O(h) \tag{15}$$

and

$$\lambda_i - \mu_i = -\{g, f\}^{-1}b_i\{V, f\}^T + O(h) \tag{16}$$

where we used that $\{g,T\} = f = 0$.

Since $X_{\hat{H}_i}$ has the first integrals $g = f = 0$, its time-t-flow $\exp(tX_{\hat{H}_i})$ on $\mathcal{M}$ is uniquely determined by the restriction of $\hat{H}_i(h)$ to $\mathcal{M}$ [2]. (Or, in other words, $X_{\hat{H}_i}(h)$ is a vector field on $\mathcal{M}$ and therefore the perturbed vector field in the sense of our Theorem 1.) Thus, using (15) and (16) and replacing the Lie-Poisson bracket $\{.,.\}$ by the Lie-Poisson bracket $\{.,.\}_{\mathcal{M}}$ on $\mathcal{M}$, we obtain after a bit of tedious but straightforward algebra the simplified

$$\begin{aligned}\hat{H}_i(h) \quad = \quad & b_iV + B_iT + \frac{h}{2}b_iB_i\{V,T\}_{\mathcal{M}} + \frac{h^2}{12}(b_iB_i^2\{T,\{T,V\}_{\mathcal{M}}\}_{\mathcal{M}} + b_i^2B_i\{V,\{V,T\}_{\mathcal{M}}\}_{\mathcal{M}} + \\ & b_iB_i^2E_1 + b_i^2B_iE_2 + B_i^3E_3) + O(h^3)\end{aligned}$$

with

$$E_1 = -2\{T,f\}\{g,f\}^{-1}\{V,f\}^T\,,$$

$$E_2 = 4\{V,f\}\{g,f\}^{-1}\{V,f\}^T\,,$$

and

$$E_3 = -\frac{1}{2}\{T,f\}\{g,f\}^{-1}\{T,f\}^T.$$

Let $X_{\hat{H}_i}(h) = \{\mathrm{id},\hat{H}_i(h)\}_{\mathcal{M}}$ denote the vector field on $\mathcal{M}$ corresponding to the Hamiltonian $\hat{H}_i(h)$, then (14) can be rewritten as

$$\Psi_h \quad = \quad \prod_{i=1}^{s}\exp(hX_{\hat{H}_i}(h))$$

Application of the BCH formula of Corollary 1 to this product results in

$$\Psi_h \quad = \quad \exp(hX_{\hat{H}}(h))$$

with

$$\begin{aligned}\hat{H}(h) \quad = \quad & \beta_1V + \beta_2T + \frac{h}{2}\beta_3\{V,T\}_{\mathcal{M}} + \\ & + \frac{h^2}{12}(\beta_4\{V,\{V,T\}_{\mathcal{M}}\}_{\mathcal{M}} + \beta_5\{T,\{,T,V\}_{\mathcal{M}}\}_{\mathcal{M}} + \beta_6E_1 + \beta_7E_2 + \beta_8E_3) + O(h^3)\,.\end{aligned}$$

Comparison with (7) yields $\beta_1 = \alpha_1$, $\beta_2 = \alpha_2$, $\beta_3 = \alpha_3$, $\beta_4 = \alpha_4$, $\beta_5 = \alpha_5$, and the three additional equations

$$\beta_6 \quad = \quad \sum_i b_iB_i^2\,,$$

$$\beta_7 \quad = \quad \sum_i b_i^2B_i\,,$$

and

$$\beta_8 \quad = \quad \sum_i B_i^3\,.$$

Summarizing our discussion so far, we obtain

Proposition 1. Any first or second order PRK method (3) generalizes to the constrained scheme (4) without order reduction. For higher-order methods we obtain, in general, additional order conditions. For example, to achieve order three, one has to satisfy the additional conditions $\beta_6 = \beta_7 = \beta_8 = 0$.

6 Higher-order PRK methods by composition of first and second order PRK methods

To avoid the growing complexity of additional order conditions, it seems necessary to restrict oneself to semi-explicit PRK methods based on the composition of first and/or second order PRK methods. Using the backward error analysis given in Section 4, it follows immediately that the order results for unconstrained systems apply also to methods based on the composition of constrained first and/or second order symplectic integrators. One can, for example, construct higher-order methods by a proper composition of the first order semi-explicit PRK method ϕ_h^+ defined by

$$\begin{aligned} q_{n+1} &= q_n + h\nabla T(p_{n+1})\,, \\ p_{n+1} &= p_n - h\nabla V(q_n) - hG(q_n)^T\lambda\,, \\ 0 &= g(q_{n+1})\,, \end{aligned} \tag{17}$$

and its adjoint PRK method ϕ_h^-:

$$\begin{aligned} q_{n+1} &= q_n + h\nabla T(\bar{p}_n)\,, \\ \bar{p}_n &= p_n - hG(q_n)^T\lambda\,, \\ 0 &= g(q_{n+1})\,, \\ \\ p_{n+1} &= \bar{p}_n - h\nabla V(q_{n+1})\,. \end{aligned} \tag{18}$$

Theorem 3. Applying formula (8) with weights $c_1 = \tilde{b}_1$, $d_1 = \tilde{B}_1 - \tilde{b}_1$,

$$c_i = \tilde{b}_i - d_{i-1}\,,$$

and

$$d_i = \tilde{B}_i - c_i$$

for $i = 2,\ldots,s$, one can transform any s-stage PRK method (3) of order p into a constrained PRK method (4) of the same order. The number of stages is $2s-1$. The coefficients (b_i) and (B_i) are given by

$$\begin{array}{ll} b_{2i-1} = \tilde{b}_i, & b_{2i} = 0, \\ B_{2i-1} = c_i, & B_{2i} = d_i, \end{array} \qquad i = 1,2,\ldots,s\,.$$

Thus, although the method (4) has $2s-1$ stages, it only requires s evaluations of $\nabla V(q)$. In each step we have to solve $2s-1$ nonlinear systems of equations in the λ_i's.

Proof. We have to show first that (18) is indeed the adjoint method to (17). As done before in Section 5, we formally add to (17) a projection step onto $f = 0$ which yields

$$\begin{aligned} q_{n+1} &= q_n + h\nabla T(\bar{p}_{n+1})\,, \\ \bar{p}_{n+1} &= p_n - h\nabla V(q_n) - hG(q_n)^T\lambda\,, \\ 0 &= g(q_{n+1})\,, \\ \\ p_{n+1} &= \bar{p}_{n+1} - hG(q_{n+1})^T\mu\,, \\ 0 &= f(q_{n+1},p_{n+1})\,. \end{aligned}$$

In terms of the mapping Ψ_h, this does not change anything. Now, $(\phi^+_{-h})^{-1}$ results formally in

$$\begin{aligned} q_{n+1} &= q_n + h\nabla T(\bar{p}_n)\,, \\ p_{n+1} &= \bar{p}_n - h\nabla V(q_{n+1}) - hG(q_{n+1})^T\lambda\,, \\ 0 &= g(q_n)\,, \\ \\ \bar{p}_n &= p_n - hG(q_n)^T\mu\,, \\ 0 &= f(q_n,p_n)\,. \end{aligned} \tag{19}$$

Noting that ϕ^+_h is a map from $\mathcal{M}$ to $\mathcal{M}$ and rearranging equations, we obtain

$$\begin{aligned} \bar{p}_n &= p_n - hG(q_n)^T\lambda\,, \\ q_{n+1} &= q_n + h\nabla T(\bar{p}_n)\,, \\ 0 &= g(q_{n+1})\,, \\ \\ p_{n+1} &= \bar{p}_n - h\nabla V(q_{n+1}) - hG(q_{n+1})^T\mu\,, \\ 0 &= f(q_{n+1},p_{n+1}) \end{aligned} \tag{20}$$

where we also relabeled μ by λ, λ by μ respectively. We finally remove the projection step onto $f = 0$ again and obtain (18) as desired.

The order results follow now immediately from our Corollary 2 and Theorem 2 in [9].

Since each c_i, d_i respectively, corresponds to one stage in a PRK method, we obtain the number of stages for our semi-explicit PRK method (4) as stated above. Note that $d_s = 0$.

The proper coefficients (b_i) and (B_i) for the PRK method (4) follow from

$$\begin{aligned} \phi^-_{d_ih}\cdot\phi^+_{c_ih} &= \exp(d_ihX_V)\cdot\exp(d_ihX_T)\cdot\exp(hX_{g^T\lambda_{i,2}})\cdot \\ &\quad\cdot\exp(c_ihX_T)\cdot\exp(hX_{g^T\lambda_{i,1}})\cdot\exp(c_ihX_V) \end{aligned} \tag{21}$$

and

$$\exp(c_{i+1}hX_V)\cdot\exp(d_ihX_V) = \exp(\tilde{b}_ihX_V)\,.$$

□

Compared to (3), we have now to solve up to $2s-1$ systems of equations in λ per integration step while the number of force evaluations $\nabla V(q)$ is still s, in general. There is one particular case where a s-stage PRK method (4) does not suffer from order reduction when $(b_i) = (\tilde{b}_i)$ and $(B_i) = (\tilde{B}_i)$. This is the subject of

Theorem 4. Consider a s-stage PRK method (3) of order p with coefficients $(\hat{b}_i)$ and $(\hat{B}_i)$ satisfying $\hat{B}_1 = 2\hat{b}_1$ and

$$\hat{B}_i = 2(\hat{b}_i - \hat{b}_{i-1})$$

for $i = 2,\ldots,s-1$. Then the corresponding constrained PRK method (4) with $(b_i) = (\hat{b}_i)$ and $(B_i) = (\hat{B}_i)$ has the same order as (3).

Proof: For unconstrained systems, the corresponding PRK method (3) is equivalent to the composition method

$$\Psi_h = \prod_{i=1}^{s-1}\psi_{B_ih}$$

where

$$\psi_h = \exp(\frac{h}{2}X_V) \cdot \exp(hX_T) \cdot \exp(\frac{h}{2}X_V) \,. \tag{22}$$

Now the order conditions for these composition schemes do not depend on the particular choice of the integrator ψ_h as long as ψ_h is symmetric and of second order [17]. It follows from the discussion in Section 4 that this result also applies to constrained systems. Thus we can replace (22) by the symmetric second-order method

$$\psi_h = \exp(hX_{g^T\mu}) \cdot \exp(\frac{h}{2}X_V) \cdot \exp(hX_T) \cdot \exp(\frac{h}{2}X_V) \cdot \exp(hX_{g^T\lambda}) \tag{23}$$

without introducing order reduction in the corresponding PRK method. Note that the method (23) is equivalent to the SHAKE or RATTLE method used in molecular dynamics [1], [8]. □

The schemes considered in Theorem 4 do not have redundant stages compared to the PRK methods for unconstrained systems. Therefore, they seem computationally more attractive than the, although more general, PRK methods discussed in Theorem 3. However, this depends on whether the evaluation of the force field or the solution of the nonlinear equations in the λ_i's are the main source of the computational expenses. We will come back to this in more detail in the following section.

7 Numerical experiment

In this section we compare several symmetric second, fourth, and sixth order methods based on an effective error constant e_{eff} defined by

$$e_{\text{eff}} = \beta^p \frac{\max|\Delta H|}{h^p}$$

where h is the step-size, β a scaling factor representing the computational expenses per integration step, and ΔH is the error in the Hamiltonian along the numerically computed solution.

Remark. The scaling factor β^p allows one to compare methods of the same order based on their local error in energy

$$e_{\text{loc}} = \frac{\max|\Delta H|}{h^p}$$

and the amount of computational work needed per integration step. For example, let us assume that we have two methods of order $p = 2$; one with 2 stages and one with 4 stages. In this case, one could, for example, set $\beta = 1$ for the first method and $\beta = 2$ for the second method (important is only the ratio between the two β's which should be equal to a half). Let us also assume that the local error of the first method is twice the local error of the second method. Thus one could say that the second method is better than the first method. However, compared on the basis of an equal amount of computational work, the first method has an effective step-size that is half the step-size of the second method. This is reflected by the scaling factor β^2 in the effective error constant e_{eff}. Consequently, in this example, the first method is more efficient than the second method.

Here we evaluated e_{loc} by integrating the mathematical double pendulum moving in a gravitational field; i.e.,

$$\begin{aligned} q_1' &= p_1\,, & p_1' &= -2\lambda_1 q_1 - 2\lambda_2(q_1 - q_3)\,, \\ q_2' &= p_2\,, & p_2' &= -g - 2\lambda_1 q_2 - 2\lambda_2(q_2 - q_4)\,, \\ q_3' &= p_3\,, & p_3' &= -2\lambda_2(q_3 - q_1)\,, \\ q_4' &= p_4\,, & p_4' &= -g - 2\lambda_2(q_4 - q_2)\,, \end{aligned}$$

$$\begin{aligned} 0 &= q_1^2 + q_2^2 - L_1^2\,, \\ 0 &= (q_1 - q_3)^2 + (q_2 - q_4)^2 - L_2^2 \end{aligned}$$

with $g = 10$, $L_1 = L_2 = 1.0$.

For a symmetric s-stage method falling under Theorem 4 (and abbreviated from now on by $SPRK$), we have $s-1$ evaluations of ∇V and $s-1$ nonlinear equations in the λ_i's. For that reason we set $\beta = (s-1)/10$ for SPRK methods. (To make sure that $\beta < 1$ for all the methods considered in this section, we divide $(s-1)$ by ten.) The situation differs for symmetric methods falling under Theorem 3 (and abbreviated from now on by FPRK). Here we have $(s-1)/2$ evaluations of ∇V and $s-1$ nonlinear equations in the λ_i's. Thus, if the evaluation of ∇V is much more expensive than the solution of the nonlinear equations in λ, then we take $\beta = (s-1)/20$ and denote the corresponding effective error constant by e_{eff}. In the opposite case, we have $\beta = (s-1)/10$ and we denote the corresponding effective error constant by e_{eff}^0. (Note that s denotes here the number of stages of the constrained PRK method (4).) The effective error constants of several methods can be found in Tab. 1. The coefficients (b_i) and (B_i) for each of these methods are given in Tab. 2 where we dropped those coefficients that can be obtained from Theorem 3, Theorem 4 respectively, or are determined by the fact that all the methods considered here are symmetric. The coefficients for the corresponding unconstrained PRK methods can be found in [9].

discretization scheme	stages	order	e_{eff}	e_{eff}^0
PRK	2	2	0.37	–
PRK	3	2	0.10	–
FPRK	9	4	1.13	18.2
FPRK	11	4	0.50	8.1
SPRK	4	4	11.2	–
SPRK	6	4	3.90	–
SPRK	8	6	82.2	–
SPRK	10	6	34.1	–

Table 1: Effective error constants

Apart from the fact that we use a different effective error constant than [9], our results for e_{eff} are similar to the error constants of explicit PRK methods applied to unconstrained systems as discussed in [9]. Note that for the fourth order methods the choice of an optimal method depends on whether the force evaluations or the nonlinear constraint equations dominate the computations. In the first case the fourth order FPRK method with 11 stages (5 force evaluations) should be used while in the second case the fourth order SPRK with 6 stages (5 nonlinear systems of equations) seems optimal.

We also implemented the fourth order FPRK methods with the coefficients (b_i) and (B_i) chosen

discretization scheme	stages	order	coefficients (b_i) and (B_i)
PRK	2	2	$b_1 = 0.5$
PRK	3	2	$b_1 = 0.1932$, $B_1 = 0.5$
FPRK	9	4	$b_1 = \frac{642+\sqrt{471}}{3924}$, $B_2 = 6/11 - b_1$, $b_3 = \frac{121}{3924}(12 - \sqrt{471})$, $B_4 = 0.5 - b_1 - b_3$
FPRK	11	4	$b_1 = \frac{14-\sqrt{19}}{108}$, $B_2 = 2/5 - b_1$, $b_3 = \frac{20-7\sqrt{19}}{108}$, $B_4 = 3/10 - b_1 - b_3$, $b_5 = 0.5 - b_1 - b_3$, $B_5 = 3/10$
SPRK	4	4	$B_1 = (2 - \sqrt[3]{2})^{-1}$
SPRK	6	4	$B_1 = 0.28$, $B_2 = 0.625466428467$
SPRK	8	6	$B_1 = 0.784513610477$, $B_2 = 0.235573213359$ $B_3 = -1.177679984178$
SPRK	10	6	$B_1 = 0.1867$, $B_2 = 0.555497023712$ $B_3 = 0.129466948913$, $B_4 = -0.843265623387$

Table 2: Coefficients for constrained PRK methods

as for the unconstrained method (3) and observed that the order of the constrained schemes dropped to two.

Acknowledgments. This work was partially supported by the Swiss Science Foundation and was started while the author was visiting Ernst Hairer and Gerhard Wanner in Geneva.

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., *Dynamical Systems III*, Springer-Verlag, NY, 1988.

[3] Benettin, G., Giorgilli, A., On the Hamiltonian Interpolation of Near to the Identity Symplectic Mappings with Application to Symplectic Integration Algorithms, *J. Statist. Phys.*, 74(1994), 1117–1143.

[4] Hairer, E., Backward Analysis of Numerical Integrators and Symplectic Methods, *Annals of Numerical Mathematics* 1(1994), 107–132.

[5] Hairer, E., Nørsett, S.P., Wanner, G., *Solving Ordinary Differential Equations, I*, second edition, Springer-Verlag, 1993.

[6] Jay, L., Symplectic Partitioned Runge-Kutta Methods for Constrained Hamiltonian Systems, *SIAM J. Num. Anal.*, to appear.

[7] Leimkuhler, B. and Reich, S., Symplectic integration of constrained Hamiltonian systems, *Math. Comput.*, 63(1994), 589–605.

[8] Leimkuhler, B. and Skeel, R.D., Symplectic Numerical Integrators in Constrained Hamiltonian Systems, *J. Comp. Phys.*, 112(1994), 117–125.

[9] Mc Lachlan, R.I., On the numerical integration of ordinary differential equations by symmetric composition methods, *SIAM J. Sci. Comput.*, 16(1995), 151–168.

[10] McLachlan, R.I., Scovel, C., Equivariant Constrained Symplectic Integration, *J. Nonlinear Science*, to appear.

[11] Okunbor, D. and Skeel, R.D., Explicit canonical methods for Hamiltonian systems, *Math. Comput.*, 59(1992), 439–455.

[12] Ostermann, A., A class of half-explicit Runge-Kutta methods for differential-algebraic systems of index 3, *APNUM*, 13(1993), 165–179.

[13] Reich, S. Symplectic integration of constrained Hamiltonian systems by Runge-Kutta methods, Techn. Report 93-13, Dept. of Comput. Sci., University of British Columbia, 1993.

[14] Reich, S. Numerical Integration of the Generalized Euler Equations, Techn. Report 93-20, Dept. of Comput. Science, University of British Columbia, 1993.

[15] Reich, S., Symplectic Integration of Constrained Hamiltonian Systems by Composition Methods, *SIAM J. Numer. Anal.*, to appear.

[16] Varadarajan, V.S., *Lie Groups, Lie Algebras, and Their Representation*, Prentice-Hall, Englewood Cliffs, 1974.

[17] Yoshida, H., Construction of higher order symplectic integrators, *Phys. Lett. A*, 150(1990), 262–269.