Order Barriers for Symplectic Multi-Value Methods

E. Hairer[*] P. Leone[*]

July 4, 1997

Abstract We study the question whether multistep methods or general multi-value methods can be symplectic. Already a precise understanding of symplecticity of such methods is a nontrivial task, because they advance the numerical solution in a higher dimensional space (in contrast to one-step methods).

The essential ingredient for the present study is the formal existence of an invariant manifold, on which the multi-value method is equivalent to a one-step method. Assuming this underlying one-step method to be symplectic, we prove that the order of the multi-value method has to be at least twice its stage order. As special cases we obtain: multistep methods can never be symplectic; the only symplectic one-leg method is the implicit mid-point rule, and the Gauss methods are the only symplectic Runge-Kutta collocation methods.

1 Introduction

We consider the numerical solution of Hamiltonian systems

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

where $p, q \in \mathbb{R}^d$. It is well-known that the use of symplectic one-step methods has favourable properties for long-time integrations (no secular terms in the error of the energy integral, linear error growth in the angle variables instead of quadratic growth, correct qualitative behaviour) as long as they are applied with constant step sizes. Since multistep methods or general multi-value methods, when used in constant step size mode, need in general less function evaluations than one-step methods, it is quite natural to investigate the symplecticity of such methods.

The first difficulty that arises is that of a suitable definition of symplecticity of multistep or, more general, multi-value methods. Recall that the exact flow of (1.1) is a symplectic transformation in the phase space $\mathbb{R}^{2d}$, i.e., it preserves the differential

[*]E-mail: Ernst.Hairer@math.unige.ch and Pierre.Leone@math.unige.ch
URL: http://www.unige.ch/math/folks/hairer/

2-form $\omega = dp \wedge dq = \sum_{i=1}^{d} dp^i \wedge dq^i$ (here, p^i, q^i denote the ith components of p, q). It is straightforward to define the symplecticity of one-step methods $y_{n+1} = \Phi_h(y_n)$, because they advance the numerical solution in the same space. However, for multi-value methods a satisfactory definition of symplecticity is less obvious.

In this article we consider three classes of numerical methods: multistep methods, one-leg methods, and general linear methods. For the system (1.1), written in the form $y' = f(y)$, *linear multistep methods* are

$$\sum_{j=0}^{k} \alpha_j y_{n+j} = h \sum_{j=0}^{k} \beta_j f(y_{n+j}), \tag{1.2}$$

and the corresponding *one-leg methods* are

$$\sum_{j=0}^{k} \alpha_j y_{n+j} = h f\Big(\sum_{j=0}^{k} \beta_j y_{n+j} \Big). \tag{1.3}$$

Throughout this paper we assume the methods to be stable, consistent, and with normalization $\sum_{j=0}^{k} \beta_j = 1$. We also consider *general linear methods*

$$\begin{aligned} u_i^{(n+1)} &= \sum_{j=1}^{k} a_{ij} u_j^{(n)} + h \sum_{j=1}^{s} b_{ij} f(v_j^{(n)}) \\ v_i^{(n)} &= \sum_{j=1}^{k} \widetilde{a}_{ij} u_j^{(n)} + h \sum_{j=1}^{s} \widetilde{b}_{ij} f(v_j^{(n)}). \end{aligned} \tag{1.4}$$

All these methods have in common that they advance the numerical solution in the higher-dimensional space $(\mathbb{R}^{2d})^k$.

The outline of this article is as follows. We start by considering linear problems (Sect. 2) and we deduce some necessary conditions for symplecticity. In Sect. 3 we show the formal existence of an underlying one-step method for each of the above methods. This allows us to give a natural definition of symplecticity. Based on the characterization of symplectic B-series of Calvo and Sanz-Serna [2], we derive order barriers (Sect. 4) for symplectic multi-value methods. Further results are presented in Sect. 5, which indicate that symplectic multi-value methods do probably not exist unless they are equivalent to one-step methods.

The presented order barriers remain all valid for partitioned methods, where the variables p and q are discretized by different multistep, one-leg, or general linear methods. Despite of the negative results presented in this paper, partitioned multi-value methods are interesting for long-time integration of Hamiltonian systems, as demonstrated by Quinlan and Tremaine [9]. However, instead of focusing on symplectic methods one has to use special classes of symmetric methods.

2 The Linear Case

For the harmonic oscillator $H(p, q) = \frac{1}{2}(p^2 + q^2)$ the differential equation (1.1) becomes $y' = iy$, if we put $y = p + iq$. The application of a multistep (or one-leg) method to

$y' = \lambda y$ yields a linear recurrence relation for the $\{y_n\}$, whose solution is given by

$$y_n = c_1(z)\zeta_1^n(z) + c_2(z)\zeta_2^n(z) + \ldots + c_k(z)\zeta_k^n(z), \tag{2.1}$$

where $z = \lambda h$ and $\zeta_1(z), \ldots, \zeta_k(z)$ are the roots of

$$\sum_{j=0}^{k} (\alpha_j - z\beta_j)\zeta^j = 0. \tag{2.2}$$

The principal root $\zeta_1(z)$ is the analytic continuation of $\zeta_1(0) = 1$ and approximates the exponential function as $\zeta_1(z) = e^z + \mathcal{O}(z^{p+1})$[1]. The coefficent $c_1(z)$ satisfies $c_1(z) = y_0 + \mathcal{O}(z^p)$, if the starting values approximate the exact solution to this order. Furthermore, $c_j(z) = \mathcal{O}(z^p)$ for $j = 2, \ldots, k$. Whenever the parasitic roots $\zeta_2(z), \ldots, \zeta_k(z)$ are bounded by 1, the numerical solution (2.1) is well approximated by $\zeta_1^n(z)y_0$. Therefore, it is natural to require that the mapping $y_0 \mapsto \zeta_1(z)y_0$ (the underlying one-step method) be symplectic. For the harmonic oscillator (one degree of freedom) symplecticity is equivalent to area preservation, which again is equivalent to $|\zeta_1(ih)| = 1$. Since $\zeta_1(z)$ is analytic in a neighbourhood of $z = 0$, this condition becomes

$$\zeta_1(z)\zeta_1(-z) = 1. \tag{2.3}$$

For general linear methods (1.4), applied to $y' = \lambda y$, we get

$$u^{(n+1)} = S(z)u^{(n)}, \qquad S(z) = A + zB(I - z\widetilde{B})^{-1}\widetilde{A}. \tag{2.4}$$

This recursion can be solved and yields for $u^{(n)}$ an approximation as in Eq. (2.1), where $\zeta_j(z)$ are the eigenvalues of $S(z)$ and $c_j(z)$ the corresponding eigenvectors. Therefore, $\zeta_j(z)$ are the zeros of the characteristic polynomial

$$Q(z,\zeta) := \det(I - z\widetilde{B})\,\det(\zeta I - S(z)) = q_k(z)\zeta^k + q_{k-1}(z)\zeta^{k-1} + \ldots + q_0(z). \tag{2.5}$$

The polynomials $q_j(z)$ have degree at most s (see [7], Lemma V.4.12).

Lemma 2.1 *For irreducible linear multistep methods (or one-leg methods) the symplecticity for linear problems (i.e., condition (2.3)) is equivalent to the symmetry of the method, i.e.,*

$$\alpha_j = -\alpha_{k-j}, \qquad \beta_j = \beta_{k-j}, \qquad j = 0, \ldots, k. \tag{2.6}$$

In the case of stable general linear methods with irreducible polynomial $Q(z,\zeta)$ the symplecticity for linear problems is equivalent to

$$q_j(z) = -q_{k-j}(-z), \qquad j = 0, \ldots, k. \tag{2.7}$$

Hence, the eigenvalues of the matrix A lie all on the unit circle.

---

[1] For historical reasons we use the same letter p for the variable in Eq. (1.1) and for the order of a numerical method. Another historical incident is the simultaneous use of q for the Lagrangian coordinates as well as for the stage order and also for the coefficients in (2.5). We appologize.

Proof. Replacing in Eq. (2.5) z by $-z$ and ζ by $1/\zeta$, we see that the condition (2.3) implies that $\zeta_1(z) = 1/\zeta_1(-z)$ is also a root of the equation

$$q_0(-z)\zeta^k + q_1(-z)\zeta^{k-1} + \ldots + q_k(-z) = 0. \tag{2.8}$$

The solution $\zeta_1(z)$ of $Q(z,\zeta) = 0$, defined on a neighbourhood of the origin, uniquely determines all other roots by analytic continuation. By the irreducibility of $Q(z,\zeta)$ the polynomials (2.5) and (2.8) are therefore equal up to a complex constant. Using the fact that $Q(0,1) = 0$ and $\frac{\partial Q}{\partial \zeta}(0,1) \neq 0$, this constant is seen to be -1. □

This lemma shows that for the study of symplectic multi-value methods one has to consider weakly stable methods.

3 The Underlying One-Step Method

For nonlinear differential equations the definition of symplecticity for multi-value methods is less evident. One has essentially two possibilities: either

- extend the differential form $dp \wedge dq$ to the space $(\mathbb{R}^{2d})^k$; or
- restrict the flow of the numerical method to a submanifold of dimension $2d$.

The first approach is that of Eirola and Sanz-Serna [4], where it is shown that symmetric one-leg methods preserve the differential form $\sum_{i,j=1}^{k} g_{ij} dp_{n+i} \wedge dq_{n+j}$ for a suitable matrix G. Since the exact flow of (1.1), extended componentwise to the product space $(\mathbb{R}^{2d})^k$, does in general not preserve such a differential form (see Lemma 5.1 below), we follow the second approach.

Kirchgraber [8] proved that the mapping

$$\Psi_h : (y_n, y_{n+1}, \ldots, y_{n+k-1})^T \mapsto (y_{n+1}, y_{n+2}, \ldots, y_{n+k})^T,$$

defined by a strictly stable multistep formula, possesses an attractive invariant manifold on which the multistep method is equivalent to a one-step method $\Phi_h(y)$. The manifold is given by

$$\mathcal{M} = \left\{ (y, \Phi_h(y), \Phi_h^2(y), \ldots, \Phi_h^{k-1}(y))^T \;\middle|\; y \in \mathbb{R}^{2d} \right\}$$

so that, whenever the starting values satisfy $y_j = \Phi_h^j(y_0)$ for $j = 0, 1, \ldots, k-1$, then the numerical solution of the multistep method and that of the underlying one-step method $\Phi_h(y)$ are identical. A generalization of this result to general linear methods (1.4) can be found in Stoffer [10].

It is tempting to define a multi-value method to be symplectic, if its underlying one-step method is symplectic. However, because of Lemma 2.1 we would need Kirchgraber's result also for weakly stable methods. The proofs of [8] and [10] are based on invariant manifold theory and on fixed point arguments and cannot be extended easily to cover weakly stable methods. We therefore try to construct explicitly the one-step method as a series in powers of h.

Inspired by the Taylor series expansion of one-step methods (like Runge-Kutta methods) we are looking for a series (called B-series) of the form

$$\begin{aligned}\Phi_h(y) &= y + ha(\bullet)f(y) + \frac{h^2}{2!}a([\bullet])f'(y)f(y) + \dots \\ &= y + \sum_{t\in T}\frac{h^{\rho(t)}}{\rho(t)!}\alpha(t)a(t)F(t)(y).\end{aligned} \tag{3.1}$$

Here, we use the notation of [6], Sect. II.2: $T = \{\bullet, [\bullet], \dots\}$ denotes the set of rooted trees, $\rho(t)$ the number of vertices of a tree $t \in T$ (also called the order of t), and the integer $\alpha(t)$ counts the number of monotonic labellings of t. The functions $F(t)(y)$ are called elementary differentials of $f(y)$ and are recursively defined by $F(\bullet)(y) = f(y)$ and

$$F(t)(y) = f^{(m)}(y)\cdot\big(F(t_1)(y), \dots, F(t_m)(y)\big) \qquad \text{for} \quad t = [t_1, \dots, t_m].$$

The tree $t = [t_1, \dots, t_m]$ is obtained from $t_1, \dots, t_m$ by adding a new vertex (the root of t) and by connecting it with the roots of $t_1, \dots, t_m$. The real coefficients $a(t)$ (for $t \in T$) have to be determined in such a way that the values $y_j = \Phi_h^j(y_0)$ satisfy the multistep formula (1.2).

Theorem 3.1 *If the polynomial $\rho(\zeta) = \sum_{j=0}^k \alpha_j\zeta^j$ of the multistep formula (1.2) satisfies $\rho(1) = 0$ and $\rho'(1) \neq 0$, then there exist uniquely determined coefficients $a(t)$ such that the series $\Phi_h(y)$ of (3.1) satisfies*

$$\sum_{j=0}^{k}\alpha_j\Phi_h^j(y) = h\sum_{j=0}^{k}\beta_j f(\Phi_h^j(y)) \tag{3.2}$$

in the sense of formal power series.

Proof. We denote the series (3.1) by $B(a, y)$, and consider a second series of the same structure but with $a(t)$ replaced by $b(t)$. It is known (Theorem II.12.6 of [6]) that the composition of these series is again a series of the same type

$$B(b, B(a, y)) = B(ab, y),$$

where $(ab)(t) = a(t) + b(t) + \dots$. The dots indicate expressions which only depend on trees with less than $\rho(t)$ vertices. Consequently, the jth iterate of $\Phi_h(y_0) = B(a, y_0)$ satisfies $\Phi_h^j(y_0) = B(a^j, y_0)$ with

$$a^j(t) = ja(t) + \dots \; . \tag{3.3}$$

Next we apply Corollary II.12.7 of [6] and obtain that

$$hf\big(\Phi_h(y_0)\big) = \sum_{t\in T}\frac{h^{\rho(t)}}{\rho(t)!}\alpha(t)a'(t)F(t)(y_0),$$

where $a'(\bullet) = 1$ and $a'(t) = \rho(t)a(t_1) \cdot \ldots \cdot a(t_m)$ for $t = [t_1, \ldots, t_m]$. Inserting all these expressions into (3.2) and comparing the coefficients we see that $y_j = \Phi_h^j(y_0)$ satisfies formally the multistep method if

$$\sum_{j=0}^{k} \alpha_j a^j(t) = \sum_{j=0}^{k} \beta_j (a^j)'(t) \qquad \text{for all} \quad t \in T. \tag{3.4}$$

By (3.3) this is an equation of the type $(\sum_{j=0}^{k} j\alpha_j)a(t) = \ldots$, where the right-hand expression only depends on trees with less than $\rho(t)$ vertices. Since $\sum_j j\alpha_j \neq 0$, this equation uniquely defines the coefficients $a(t)$. □

It is natural to ask, whether the series (3.1) with the coefficients $a(t)$ from the preceding theorem converges. For linear problems this function corresponds to $\zeta_1(\lambda h)y_0$ of Sect. 2. Since $\zeta_1(z)$ is analytic in a neighbourhood of 0, we have convergence. This convergence, however, is exceptional as can be seen from the following example.

Example Consider a consistent two-step method

$$\alpha_2 y_{n+2} + \alpha_1 y_{n+1} + \alpha_0 y_n = h\Big(\beta_2 f(y_{n+2}) + \beta_1 f(y_{n+1}) + \beta_0 f(y_n)\Big),$$

and apply it to the simple system $y' = f(x)$, $x' = 1$. The y-component of the series (3.1) then takes the form

$$\Phi_h(x_0, y_0) = y_0 + \sum_{\rho \geq 1} \frac{h^\rho}{\rho!}\, a_\rho\, f^{(\rho-1)}(x_0). \tag{3.5}$$

Putting $f(x) = e^x$ yields

$$A(\zeta) = \sum_{\rho \geq 1} \frac{a_\rho}{\rho!}\, \zeta^{\rho-1} = \frac{\beta_2 e^{2\zeta} + \beta_1 e^{\zeta} + \beta_0}{\alpha_2(1 + e^{\zeta}) + \alpha_1}.$$

for the generating function of the coefficients a_ρ. Since this function has finite poles, the radius of convergence of $A(\zeta)$ is finite. Therefore, the radius of convergence of the series (3.5) has to be zero as soon as $f^{(\rho)}(x_0)$ behaves like $\rho!\,\mu\,\kappa^\rho$ (what is typically the case for analytic functions). Independent of the fact whether the method is strictly stable or not, the series (3.5) usually does not converge.

Remark It can be shown that the coefficients $a(t)$ of Theorem 3.1 satisfy an estimate

$$|a(t)| \leq \rho(t)!\,\mu\,\kappa^{\rho(t)-1} \tag{3.6}$$

with suitably chosen constants μ and κ. This is precisely the estimate that makes a backward error analysis possible (see Lemma 10 of [5]). Details will be presented elsewhere.

3.2 *One-Leg Methods*

The underlying one-step method for one-leg methods can be obtained in exactly the same way. We only have to replace (3.4) by

$$\sum_{j=0}^{k} \alpha_j a^j(t) = \Big(\sum_{j=0}^{k} \beta_j a^j \Big)'(t) \qquad \text{for all} \quad t \in T. \tag{3.7}$$

3.3 *General Linear Methods*

In order to find the underlying one-step method of (1.4) we have to specify how the numerical solution y_n can be extracted from the vector $u^{(n)}$. Most common methods are based on the assumption that $u^{(n)}$ approximates either

$$\begin{array}{ll} \Big(y(x), y(x+h), \ldots, y(x+(k-1)h)\Big)^T & \text{(multistep case) or} \\ \Big(y(x), hy'(x), \ldots, \frac{h^{k-1}}{(k-1)!} y^{(k-1)}(x)\Big)^T & \text{(Nordsieck vector).} \end{array}$$

We therefore assume for the moment that y_n is one of the components of $u^{(n)}$ or, more general, that y_n is a linear combination of the components of $u^{(n)}$.

Throughout this article we assume that 1 is a simple eigenvalue of $A = (a_{ij})$, and that a vector z_0 exists such that

$$A z_0 = z_0, \qquad \widetilde{A} z_0 = 1\!\!1 \tag{3.8}$$

(preconsistency conditions; see [6], page 441), where $1\!\!1 = (1, \ldots, 1)^T$. We also denote the vector z_0 by $z(\emptyset)$ and its components by $z_i(\emptyset)$. We then represent the so-called starting procedure by the B-series

$$B(z_i, y) = z_i(\emptyset) y + \sum_{t \in T} \frac{h^{\rho(t)}}{\rho(t)!} \alpha(t) z_i(t) F(t)(y), \tag{3.9}$$

and we remark that Theorem II.12.6 of [6] also covers the case where $z_i(\emptyset) \neq 1$ (it is only required that $a(\emptyset) = 1$), i.e., we have $B(z_i, B(a, y)) = B(az_i, y)$, where $(az_i)(t) = z_i(\emptyset)a(t) + z_i(t) + \ldots$ (terms involving trees with less than $\rho(t)$ vertices).

Theorem 3.2 *Let η be a vector satisfying $\eta^T z_0 = 1$, and suppose that $y_n = \eta^T u^{(n)} = \sum_{j=1}^{k} \eta_j u_j^{(n)}$. Then there exist unique coefficients $a(t)$ and $z_i(t)$ for $t \in T$, such that $u_i^{(0)} = B(z_i, y_0)$, $y_1 = B(a, y_0)$ and $u_i^{(1)} = B(z_i, y_1)$ satisfy the multi-value formula (1.4) in the sense of formal power series. This means that the diagram*

$$\begin{array}{ccc} u^{(0)} & \xrightarrow{(1.4)} & u^{(1)} \\ {\scriptstyle (3.9)} \uparrow & & {\scriptstyle (3.9)} \uparrow \\ y_0 & \xrightarrow{(3.1)} & y_1 \end{array}$$

formally commutes.

Proof. We insert the series (3.9) together with $y_1 = \Phi_h(y_0)$ of (3.1) into (1.4) and compare the coefficients. This yields

$$\begin{aligned} (az)(t) &= Az(t) + Bv'(t) \\ v(t) &= \widetilde{A}z(t) + \widetilde{B}v'(t). \end{aligned} \tag{3.10}$$

Here we have used vector notation $z(t) = (z_j(t))_{j=1}^k$, $(az)(t) = ((az_j)(t))_{j=1}^k$ and similarly for $v(t)$. Observe that the second line of (3.10) serves as definition of $v(t)$ and, using $(az)(t) = z(\emptyset)a(t) + z(t) + \ldots$, the first line of (3.10) becomes

$$a(t)\, z_0 + (I - A)\, z(t) = \ldots\,, \tag{3.11}$$

where the right-hand expression only depends on $a(\tau), z(\tau)$ with $\rho(\tau) < \rho(t)$. By our assumption on A (1 is a simple eigenvalue), Eq. (3.11) together with $\eta^T z(t) = 0$ (what is equivalent to $y_n = \eta^T u^{(n)}$) uniquely determine $a(t)$ and $z(t)$ as functions of $a(\tau), z(\tau)$ with $\rho(\tau) < \rho(t)$. □

The existence of an underlying one-step method (although only in a formal sense) is interesting by its own. It allows us to give natural definitions of order, symmetry, ... for general linear methods by requiring the same property for the underlying one-step method.

3.4 Error Analysis

For strictly stable methods it follows from [8] that the invariant manifold is exponentially attractive. This implies that for arbitrary starting values $y_0, y_1, \ldots, y_{k-1}$ the numerical solution satisfies

$$y_n = \Phi_h^n(y_0^*) + \mathcal{O}(\kappa^n), \tag{3.12}$$

where $0 \le \kappa < 1$, and y_0^* is close to y_0. Therefore, the long-time behaviour of the numerical solution is completely determined by that of the underlying one-step method.

For weakly stable methods the invariant manifold cannot be expected to be attractive. All we can hope is that the numerical solution remains close to the manifold for very long times. Without going into details we shortly indicate, how the long-time behaviour of such methods can be studied. It is possible to prove that formally the numerical solution satisfies

$$y_n = \Phi_h^n(y_0^*) + \zeta_2^n z_2(nh) + \zeta_3^n z_3(nh) + \ldots\,, \tag{3.13}$$

where $\zeta_1 = 1, \zeta_2, \ldots, \zeta_k$ are the roots of the characteristic equation (2.2) (or (2.5)), and $\zeta_{k+1}, \zeta_{k+2}, \ldots$ are finite products of $\zeta_2, \ldots, \zeta_k$. Similarly to the asymptotic expansion of the global error of multistep methods, the functions $z_j(x)$ are solutions of differential equations $z_j' = \lambda_j f'(y) z_j + \ldots$, where λ_j is the growth parameter corresponding to ζ_j. If the functions $z_j(x)$ in (3.13) remain bounded and small over long time intervals, the one-step method $\Phi_h(y)$ determines the long-time behaviour of the multi-value method.

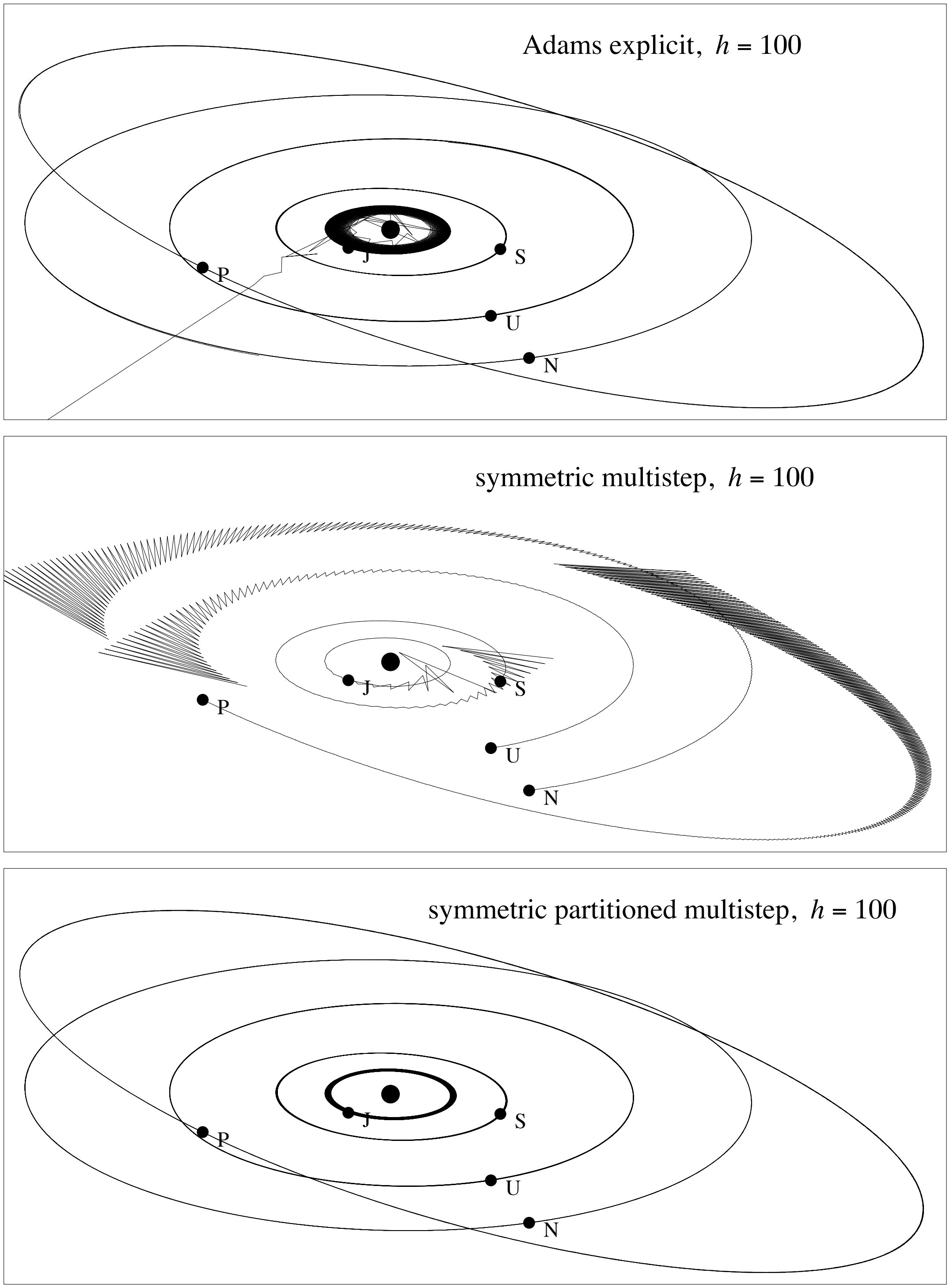

Figure 3.1: Multistep methods of order 4 applied to planetary orbits

Example The Hamiltonian for the equations of motion of the sun and the five outer planets is given by

$$H(p,q) = \frac{1}{2}\sum_{i=0}^{5} m_i^{-1} p_i^T p_i - \sum_{i=1}^{5}\sum_{j=0}^{i-1} \frac{K\, m_i\, m_j}{\|q_i - q_j\|},$$

where $q_i, p_i \in I\!R^3$, K (gravitation) and m_i (masses) are given constants.

We first apply the explicit Adams method of order 4, which is a strictly stable multistep method (first picture of Fig. 3.1). A backward error analysis of the underlying one-step method shows that the numerical solution is close to the exact solution of a perturbed differential equation with dissipative perturbations. Hence, the numerical orbits spiral inward and finally become completely wrong.

For the second experiment we take the symmetric method

$$y_{n+4} - y_n = \frac{h}{3}\Big(8f_{n+3} - 4f_{n+2} + 8f_{n+1}\Big).$$

In this case the underlying one-step method is symmetric and gives a qualitatively correct solution. However, the functions $z_j(x)$ in (3.13) grow linearly in time, which causes the increasing oscillations in the second picture of Fig. 3.1.

Finally, we apply a partitioned multistep method, where the two equations of (1.1) are discretized by two different multistep methods (with generating polynomials (ρ_1, σ_1) and (ρ_2, σ_2)). Inspired by the results of [9] and [3] we construct the methods in such a way that $\rho_1(\zeta)$ and $\rho_2(\zeta)$ have no common zero with the exception of $\zeta_1 = 1$. In this case it turns out that the functions $z_j(x)$ of (3.13) are nicely bounded and small, so that the observed numerical solution is essentially that of a symmetric one-step method (last picture of Fig. 3.1).

This example demonstrates that for a deeper understanding of the long-time behaviour of multi-value methods one has to study the boundedness of $z_j(x)$ as well as the qualitative properties of the underlying one-step method. In the rest of this paper we shall study properties of $\Phi_h(y)$.

4 Order Barriers

Calvo and Sanz-Serna [2] have given a characterization of symplectic B-series. They have shown that a transformation, defined by (3.1), is symplectic if and only if

$$\frac{a(u\circ v)}{\gamma(u\circ v)} + \frac{a(v\circ u)}{\gamma(v\circ u)} = \frac{a(u)}{\gamma(u)}\cdot\frac{a(v)}{\gamma(v)} \qquad \text{for all} \quad u, v \in T. \tag{4.1}$$

Here, the product $u \circ v$, introduced by Butcher, denotes the tree of order $\rho(u) + \rho(v)$, which is obtained by attaching the root of v via an additional branch to the root of u (which will be the root of $u \circ v$). This product is not commutative. The integer coefficient $\gamma(t)$ is defined recursively by $\gamma(\bullet) = 1$ and $\gamma(t) = \rho(t)\gamma(t_1)\cdot\ldots\cdot\gamma(t_m)$

for $t=[t_1,\ldots,t_m]$. Inspite of the fact that the series (3.1) usually does not converge, but encouraged by the considerations of Sect. 3.4, we use the following definition of symplecticity.

Definition 4.1 *A multistep method (or a general multi-value method) is called symplectic, if the coefficients $a(t)$ of the underlying one-step method satisfy (4.1).*

The main result of this section will be presented in Theorem 4.3. It contains Theorems 4.1 and 4.2 (recently also proved by Tang [11]) as special cases.

Theorem 4.1 (Multistep Methods) *Consistent multistep methods can never be symplectic.*

Proof. The coefficients $a(t)$, defined by (3.4), satisfy

$$\begin{aligned} a(t) &= 1 && \text{for } \rho(t)\le p \\ a(t) &= 1-C && \text{for } \rho(t)=p+1, \end{aligned} \tag{4.2}$$

where p is the order of the method and $C=(\sum_j \alpha_j j^{p+1}-(p+1)\sum_j \beta_j j^p)/(\sum_j j\alpha_j)$ its error constant. Indeed, (4.2) implies $a^j(t)=j^{\rho(t)}$ for $\rho(t)\le p$, $a^j(t)=j^{\rho(t)}+j(a(t)-1)$ for $\rho(t)=p+1$, $(a^j)'(t)=\rho(t)j^{\rho(t)-1}$ for $\rho(t)\le p+1$, and one can easily check that the coefficients of (4.2) satisfy (3.4) for $\rho(t)\le p+1$.

We now take two trees u and v satisfying $\rho(u)+\rho(v)=p+1$. Assuming the multistep method to be symplectic, we obtain from (4.1) that

$$(1+C)\Big(\frac{1}{\gamma(u\circ v)}+\frac{1}{\gamma(v\circ u)}\Big)=\frac{1}{\gamma(u)}\cdot\frac{1}{\gamma(v)}. \tag{4.3}$$

The exact solution of the Hamilton system, which is a series (3.1) with all coefficients equal to 1, is a symplectic transformation. Hence, by the criterion of Calvo and Sanz-Serna, we have $1/\gamma(u\circ v)+1/\gamma(v\circ u)=1/(\gamma(u)\cdot\gamma(v))$. Consequently, (4.3) implies $1+C=1$ which is not possible for a method of finite order. □

In his seminal talk at the 1975 Dundee conference G. Dahlquist pointed out the following interesting relation between multistep and one-leg methods: if $\{y_m\}$ satisfies the one-leg difference relation (1.3), then $\widehat{y}_m=\sum_{j=0}^k \beta_j y_{m+j}$ satisfies (1.2); conversely, the one-leg solution can be recovered from $\{\widehat{y}_m\}$ by a relation of the form $y_{m+\ell}=\sum_{j=0}^{k-1} a_j\widehat{y}_{m+j}-h\sum_{j=0}^{k-1} b_j f(\widehat{y}_{m+j})$ (see [7], pages 319-320). Hence, multistep and one-leg methods are expected to have the same long-time behaviour.

Theorem 4.2 (One-Leg Methods) *The highest possible order of a consistent symplectic one-leg method is two.*

Proof. In the same way as for multistep methods one finds that the coefficients $a(t)$ of a pth order one-leg method satisfy

$$\begin{aligned} a(t) &= 1 && \text{for } \rho(t)\le p \\ a(t) &= 1-C_1 && \text{for } \rho(t)=p+1 \text{ and } t=[t_1] \\ a(t) &= 1-C_2 && \text{for } \rho(t)=p+1 \text{ and } t=[t_1,\ldots,t_m] \text{ with } m\ge 2, \end{aligned} \tag{4.4}$$

where $C_1 = (\sum_j \alpha_j j^{p+1} - (p+1)\sum_j \beta_j j^p)/(\sum_j j\alpha_j)$ and $C_2 = (\sum_j \alpha_j j^{p+1} - (p+1)(\sum_j j\beta_j)^p)/(\sum_j j\alpha_j)$. Here, we have two different constants, because the local error is composed of the differentiation and the interpolation errors, which are proportional to $y^{(p+1)}(x)$ and $f'(y(x))y^{(p)}(x)$, respectively.

Suppose the one-leg method to be symplectic and of order $p \geq 3$. Considering trees u and v with $\rho(u) + \rho(v) = p+1$, $\rho(u) \geq 2$, $\rho(v) \geq 2$, we deduce $C_2 = 0$ as in the proof of Theorem 4.1. Taking $u = \bullet$ and $\rho(v) = p$ we deduce $C_1 = 0$, which gives the contradiction. □

The extension of these order barriers to general linear methods requires two definitions of order. Method (1.4) is said to have

$$\begin{array}{rcl} \text{(classical) order } p & \Longleftrightarrow & a(t) = 1 \quad \text{for} \quad \rho(t) \leq p \\ \text{stage order } q & \Longleftrightarrow & v_i(t) = c_i^{\rho(t)} \quad \text{for} \quad \rho(t) \leq q \quad \text{and for all } i. \end{array}$$

Here, $v_i(t)$ are the components of $v(t)$ defined in (3.10). We always assume $p \geq 1$, and $q \geq 1$ is satisfied by definition of the c_i.

Theorem 4.3 (General Linear Methods) *If the underlying one-step method of (1.4) is symplectic, then the order p and the stage order q are related by*

$$p \geq 2q.$$

Proof. We outline the main ideas of the proof in the following three steps.

a) *Simplification by the stage order.* Let

$$t = [t_1, \ldots, t_{m-1}, t_m] \qquad \text{and} \qquad w = [t_1, \ldots, t_{m-1}, \overbrace{\bullet, \ldots, \bullet}^{\rho(t_m)}] \tag{4.5}$$

be two trees of the same order with $\rho(t_m) \leq \min(p,q)$. Then it holds true that

$$a(t) = a(w) \qquad \text{and} \qquad z(t) = z(w). \tag{4.6}$$

This means that trees $t = [t_1, \ldots, t_{m-1}, t_m]$, where at least one of the subtrees t_j satisfies $2 \leq \rho(t_j) \leq \min(p,q)$, can be disregarded in considering the order conditions. This statement follows from the fact that $v'(t) = v'(w)$ (which is an immediate consequence of $\rho(t_m) \leq q$) and by inspecting the composition formula of B-series for $(az)(t)$. In this way, the right-hand side of (3.11) is seen to be the same for t and for w, so that also (4.6) holds.

b) *Order for the bushy trees.* Let $\tau_\alpha = [\overbrace{\bullet, \ldots, \bullet}^{\alpha-1}]$ be the bushy tree of order α. We shall prove that

$$a(\tau_\alpha) = 1 \qquad \text{for} \qquad \alpha \leq 2q. \tag{4.7}$$

Let $\alpha \leq \beta \leq q$ and assume that $a(\tau_\gamma) = 1$ for all $\gamma \leq \beta$. As in part (a) one can show that $a(\tau_\alpha \circ \tau_\beta) = a(\tau_\beta \circ \tau_\alpha) = a(\tau_{\alpha+\beta})$ (due to the special structure of the trees we do not need $\beta \leq p$). Application of the criterion (4.1) with u, v replaced by τ_α, τ_β shows that $a(\tau_\alpha) = a(\tau_\beta) = 1$ implies $a(\tau_{\alpha+\beta}) = 1$. Consequently also (4.7) holds.

c) *Order conditions for all trees of order* $\leq 2q$. Using alternatively (4.6) and (4.1), one proves by induction on $\rho(t)$ that the order condition $a(t) = 1$ has to be satisfied for all trees of order $\leq 2q$. With the help of (4.6) we first reduce a given tree t with $\rho(t) \leq 2q$ to one of type w in (4.5) until at most one tree among $\{t_1, \ldots, t_m\}$, say t_1, is of order bigger than q. We then exchange the root of t with that of t_1. By (4.1) the order condition of the new tree will be satisfied if and only if that of t is satisfied (by induction hypothesis we assume that the order conditions of lower order trees are already satisfied). We again reduce the new tree with (4.6), so that finally we arrive at a bushy tree which is satisfied by part (b). □

Remark Since multistep methods have stage order $q = p$, and one-leg methods have stage order $q = p - 1$, the statements of Theorems 4.1 and 4.2 are a special case of that of Theorem 4.3.

Aubry and Chartier [1] call a numerical one-step method *pseudo-symplectic of order* r, if the condition (4.1) is satisfied for all trees u, v with $\rho(u) + \rho(v) \leq r$. Since the order barriers of this section are all obtained by just looking at the leading term of the truncation error, all statements of this section remain valid if one replaces "symplectic" by "pseudo-symplectic of order $p + 1$".

5 Non-Existence of Symplectic Multi-Value Methods

It is an open problem whether there exist symplectic multi-value methods that are not equivalent to a one-step method. In this section we show that with the exception of the implicit mid-point rule none of the one-leg methods is symplectic. Some further considerations support the conjecture that multi-value methods cannot be symplectic unless they are equivalent to one-step methods.

5.1 One-Leg Methods

Eirola and Sanz-Serna [4] show that for every irreducible symmetric one-leg method there exists an invertible, symmetric matrix $G = (g_{ij})$ such that the numerical solution preserves the differential form

$$\sum_{i,j=1}^{k} g_{ij}\, dp_{n+i} \wedge dq_{n+j}.$$

The matrix G is defined via the relation

$$\rho(\zeta)\sigma(\omega) + \rho(\omega)\sigma(\zeta) = (\zeta\omega - 1) \sum_{i,j=1}^{k} g_{ij} \zeta^{i-1} \omega^{j-1}, \tag{5.1}$$

where $\rho(\zeta) = \sum_j \alpha_j \zeta^j, \sigma(\zeta) = \sum_j \beta_j \zeta^j$ are the generating polynomials of the one-leg method. We show in the next lemma that this property does not always imply the symplecticity (for general Hamiltonian systems) of the underlying one-step method.

Lemma 5.1 *Consider a numerical one-step method* $(p_{n+1}, q_{n+1}) = \Phi_h(p_n, q_n)$ *that is given by a formal series (3.1). If this method is symplectic (i.e., the condition (4.1) is satisfied) and if there exists a symmetric matrix* G *such that*

$$\sum_{i,j=1}^{k} g_{ij}\, dp_{n+i} \wedge dq_{n+j} = \sum_{i,j=1}^{k} g_{ij}\, dp_{n+i-1} \wedge dq_{n+j-1}, \tag{5.2}$$

then the matrix G *has to be diagonal.*

Proof. Let $a(t)$ be the coefficients of the B-series representing the one-step method, and denote by $a^i(t)$ those of the composite series $(p_{n+i}, q_{n+i}) = \Phi_h^i(p_n, q_n) = B(a^i, (p_n, q_n))$. Inserting these series into (5.2) and using the independence of "elementary products" (Calvo and Sanz-Serna [2, Sect. 5]) we see that (5.2) is equivalent to

$$\sum_{i,j=1}^{k} g_{ij}\Big(\frac{a^i(u\circ v)}{\gamma(u\circ v)} + \frac{a^j(v\circ u)}{\gamma(v\circ u)} - \frac{a^i(u)}{\gamma(u)}\frac{a^j(v)}{\gamma(v)}\Big) \\ - \sum_{i,j=1}^{k} g_{ij}\Big(\frac{a^{i-1}(u\circ v)}{\gamma(u\circ v)} + \frac{a^{j-1}(v\circ u)}{\gamma(v\circ u)} - \frac{a^{i-1}(u)}{\gamma(u)}\frac{a^{j-1}(v)}{\gamma(v)}\Big) = 0$$

for all trees $u, v \in T$. Exploiting the symmetry of G and the fact that $a(t)$ and hence also $a^i(t)$ satisfy the symplecticity relations (4.1), this condition becomes

$$\sum_{i>j} g_{ij}\Big(\frac{a^i(u)}{\gamma(u)} - \frac{a^j(u)}{\gamma(u)}\Big)\Big(\frac{a^i(v)}{\gamma(v)} - \frac{a^j(v)}{\gamma(v)}\Big) \\ - \sum_{i>j} g_{ij}\Big(\frac{a^{i-1}(u)}{\gamma(u)} - \frac{a^{j-1}(u)}{\gamma(u)}\Big)\Big(\frac{a^{i-1}(v)}{\gamma(v)} - \frac{a^{j-1}(v)}{\gamma(v)}\Big) = 0 \tag{5.3}$$

for all trees $u, v \in T$.

For the proof of the theorem it will turn out that it is enough to consider the bushy trees τ_ρ. We let $a_\rho = a(\tau_\rho)$, $a_\rho^{(i)} = a^i(\tau_\rho)$ and consider the generating functions

$$A(\zeta) = \sum_{\rho\geq 1} \frac{a_\rho}{\rho!}\,\zeta^{\rho-1}, \qquad A_i(\zeta) = \sum_{\rho\geq 1} \frac{a_\rho^{(i)}}{\rho!}\,\zeta^{\rho-1}.$$

From the composition formula for B-series we know that $a_\rho^{(i)} = a_\rho + \sum_{j=1}^{\rho}\binom{\rho}{j} a_j^{(i-1)}$ (assuming $a_1 = 1$), so that

$$A_i(\zeta) = A(\zeta) + e^{\zeta}A_{i-1}(\zeta) = (1 + e^{\zeta} + e^{2\zeta} + \ldots + e^{(i-1)\zeta})A(\zeta). \tag{5.4}$$

We now replace in Eq. (5.3) u by τ_ρ, v by τ_σ, we multiply the equation by $\zeta^{\rho-1}\omega^{\sigma-1}$ and sum over all ρ and σ. This yields

$$\sum_{i>j} g_{ij}(e^{\zeta}e^{\omega} - 1)\big(A_{i-1}(\zeta) - A_{j-1}(\zeta)\big)\big(A_{i-1}(\omega) - A_{j-1}(\omega)\big) = 0.$$

Inserting (5.4) and dividing by $(e^\zeta e^\omega - 1)A(\zeta)A(\omega)/((e^\zeta - 1)(e^\omega - 1))$, this gives the necessary condition $\sum_{i>j} g_{ij}(e^{j\zeta} - e^{i\zeta})(e^{j\omega} - e^{i\omega}) = 0$, which is equivalent to

$$\sum_{i>j} g_{ij}(j^\rho - i^\rho)(j^\sigma - i^\sigma) = 0 \qquad \text{for all} \qquad \rho \geq 1, \sigma \geq 1.$$

This is not possible unless G is a diagonal matrix. $\square$

Theorem 5.1 *The only irreducible k-step one-leg method that is symplectic (according to Definition 4.1) is given by*

$$\rho(\zeta) = (\zeta^k - 1)/k, \qquad \sigma(\zeta) = (\zeta^k + 1)/2. \tag{5.5}$$

This is the implicit mid-point rule applied with step kh.

Proof. Suppose that the one-leg method is symplectic. By Lemma 2.1 it has to be symmetric so that the existence of a symmetric matrix G, for which (5.2) holds, follows from [4]. We therefore have to study for which methods this matrix is diagonal (Lemma 5.1).

Assume that G, defined by (5.1), is diagonal. For $\omega = 0$ the right-hand expression of (5.1) is therefore constant and we have

$$\rho(\zeta)\sigma(0) + \rho(0)\sigma(\zeta) = C$$

(by our irreducibility assumption $C \neq 0$). We now replace ζ by $1/\zeta$ and multiply by ζ^k. Due to the symmetry of the method (relation (2.6)), we get

$$-\rho(\zeta)\sigma(0) + \rho(0)\sigma(\zeta) = C\zeta^k.$$

These two relations uniquely define $\rho(\zeta)$ and $\sigma(\zeta)$ which, by the normalization $\sigma(1) = 1$, are given by (5.5). $\square$

5.2 General Linear Methods

For irreducible Runge-Kutta methods one knows that symplecticity (i.e., condition (4.1)) is equivalent to

$$M = (b_i a_{ij} + b_j a_{ji} - b_i b_j)_{i,j=1}^s = 0. \tag{5.6}$$

This matrix is known from nonlinear stability theory (B-stability, algebraic stability).

In the case of general linear methods the analogous theory of Burrage and Butcher leads to the matrix

$$M = \begin{pmatrix} G - A^T G A & \widetilde{A}^T D - A^T G B \\ D\widetilde{A} - B^T G A & D\widetilde{B} + \widetilde{B}^T D - B^T G B \end{pmatrix}, \tag{5.7}$$

where G is an arbitrary symmetric matrix and D is an arbitrary diagonal matrix. Replacing scalar products by exterior products in Lemma V.9.2 of [7] we get the identity

$$\sum_{i,j=1}^{k} g_{ij}\Big(dp_i^{(1)}\wedge dq_j^{(1)} - dp_i^{(0)}\wedge dq_j^{(0)}\Big) + \sum_{i,j=1}^{k+s} m_{ij}\; dK_i\wedge dL_j = 0,$$

where $(p_i^{(n)}, q_i^{(n)})$ and (P_i, Q_i) are the components of $u_i^{(n)}$ and $v_i^{(0)}$, respectively (see Eq. (1.4)). Further, $(K_1,\ldots,K_{k+s}) = (p_1^{(0)},\ldots,p_k^{(0)}, -hH_q(P_1,Q_1),\ldots,-hH_q(P_s,Q_s))$, $(L_1,\ldots,L_{k+s}) = (q_1^{(0)},\ldots,q_k^{(0)}, hH_p(P_1,Q_1),\ldots,hH_p(P_s,Q_s))$.

Obviously, if $M=0$ then the general linear method (1.4) preserves the differential form

$$\sum_{i,j=1}^{k} g_{ij}\, dp_i^{(n)}\wedge dq_j^{(n)}. \tag{5.8}$$

But, as we have seen in the case of one-leg methods, this does not imply the symplecticity of the underlying one-step method. We have to impose restrictions on the matrix G. One possibility seems to be the following: if the numerical solution is obtained from $u^{(n)}$ by $(p_n,q_n) = y_n = \sum_{j=1}^{k}\eta_j u_j^{(n)}$ (see Theorem 3.2), then with the choice $G=\eta\eta^T$ the differential form (5.8) becomes $dp_n\wedge dq_n$, and the assumption $M=0$ would imply the symplecticity of the underlying one-step method. Unfortunately, this choice of G together with $M=0$ imply that $\eta^T A = \eta^T$, $D1\!\!1 = B^T\eta$ and $d_i\widetilde{b}_{ij} + d_j\widetilde{b}_{ji} - d_i d_j = 0$, so that the method becomes equivalent to a one-step Runge-Kutta method.

6 Final Remarks and Open Questions

A large part of the present article deals with the underlying one-step method of multi-value methods. This part is not only of interest in connection with symplectic integration, but it also allows us to give natural definitions for properties such as order, symmetry, volume-preservation, etc.

The order barriers for pseudo-symplectic methods (Sect. 4), and the results of Sect. 5 support the conjecture that multi-value methods are probably never symplectic, unless they are equivalent to one-step methods.

In contrast to the throughout negative results on symplectic multi-value methods, the article [9] shows excellent numerical results with certain multistep methods for the long-time integration in celestial mechanics. If one is interested in a theoretical justification, the conclusion of the present article is that one should

- consider symmetric (instead of symplectic) methods, for which a good long-time behaviour can be expected for reversible Hamiltonian systems;

- study the growth of the functions $z_j(x)$ in (3.13). Partitioned methods will have more favourable properties (see Cano and Sanz-Serna [3]).

References

[1] A. Aubry and P. Chartier, "Pseudo-symplectic Runge-Kutta methods", Report, Rennes (1997).

[2] M.P. Calvo and J.M. Sanz-Serna, "Canonical B-series", Numer. Math. **67**, 161-175 (1994).

[3] B. Cano and J.M. Sanz-Serna, "Error growth in the numerical integration of periodic orbits by multistep methods, with application to reversible systems", Report 1996/5, Univ. de Valladolid, Spain. To appear in IMA J. Numer. Anal.

[4] T. Eirola and J.M. Sanz-Serna, "Conservation of integrals and symplectic structure in the integration of differential equations by multistep methods", Numer. Math. **61**, 281-290 (1992).

[5] E. Hairer and Ch. Lubich, "The life-span of backward error analysis for numerical integrators", Numer. Math. **76**, 441-462 (1997).

[6] E. Hairer, S.P. Nørsett and G. Wanner, "Solving Ordinary Differential Equations I. Nonstiff Problems", 2nd revised edition, Springer Series in Comp. Math. **8**, Springer-Verlag (1993).

[7] E. Hairer and G. Wanner, "Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems", 2nd revised edition, Springer Series in Comp. Math. **14**, Springer-Verlag (1996).

[8] U. Kirchgraber, "Multi-step methods are essentially one-step methods", Numer. Math. **48**, 85-90 (1986).

[9] G.D. Quinlan and S. Tremaine, "Symmetric multistep methods for the numerical integration of planetary orbits", Astron. J. **100** (5), 1694-1700 (1990).

[10] D. Stoffer, "General linear methods: connection to one step methods and invariant curves", Numer. Math. **64**, 395-407 (1993).

[11] Y.-F. Tang, "The symplecticity of multi-step methods", Computers Math. Applic. **25** (3), 83-90 (1993).

Acknowledgement

We gratefully acknowledge Philippe Chartier, Michel Crouzeix, Alexander Ostermann and Gerhard Wanner for helpful discussions during the preparation of this article.

Ernst Hairer and Pierre Leone
Dept. de mathématiques
Université de Genève
CH-1211 Genève 24
Switzerland