UNIVERSITY OF CAMBRIDGE

Numerical Analysis Reports

Multistep methods on manifolds

Arieh Iserles

DAMTP 1997/NA13
July 1997

Department of Applied Mathematics and Theoretical Physics
Silver Street
Cambridge
England CB3 9EW

Multistep methods on manifolds

Arieh Iserles

Department of Applied Mathematics and Theoretical Physics
University of Cambridge
Silver Street, Cambridge CB3 9EW
England
email: A.Iserles@damtp.cam.ac.uk
www: http://www.damtp.cam.ac.uk/user/na/people/Arieh/

Abstract

We explore the retention of invariants by a class of time-stepping discretization methods, inclusive of multistep methods and truncated Taylor expansions. Our main result is that no such method can respect invariance on a nonlinear manifold.

1 Numerical methods, invariants and smooth manifolds

The subject matter of this paper is the retention of invariants by time-stepping discretizations of ordinary differential equations.

Let us suppose that the ordinary differential system

$$\boldsymbol{y}' = \boldsymbol{f}(t,\boldsymbol{y}), \quad t \geq 0, \qquad \boldsymbol{y}(0) = \boldsymbol{y}_0 \in \mathbb{R}^d, \tag{1.1}$$

where $\boldsymbol{f} : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^d$ is Lipschitz and $d \geq 2$, possesses an invariant. In other words, there exists a function $\rho : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, $\rho \not\equiv 0$, such that

$$\rho(\boldsymbol{y}(t), \boldsymbol{y}_0) \equiv 0, \qquad t \geq 0. \tag{1.2}$$

An alternative manner of phrasing (1.2) is that there exists an algebraic variety $\mathcal{M}(\boldsymbol{y}_0)$ such that $\boldsymbol{y}(t) \in \mathcal{M}(\boldsymbol{y}_0)$. Provided that ρ is a smooth function (as will be the case whenever $\boldsymbol{f}$ is itself smooth), the set $\mathcal{M}(\boldsymbol{y}_0)$ is a *smooth manifold.* Moreover, it is well-known that, wherever the solution of (1.1) resides on a smooth manifold $\mathcal{M} = \mathcal{M}(\boldsymbol{y}_0)$, there exists a smooth function ρ so that $\mathcal{M}(\boldsymbol{y}_0)$ can be represented in the form (1.2) (Guillemin & Pollack 1974). The terminology of manifolds and of level sets is, thus, equivalent, and this important fact underlies much of the analysis in this paper.

An invariant is said to be a *strong integral* if there exists a nonempty open neighbourhood $\mathcal{U} \subset \mathbb{R}^d$ so that ρ is defined for every $\boldsymbol{y}_0 \in \mathcal{U}$ – in many instances $\mathcal{U} = \mathbb{R}^d$. Strong integrals feature in most applications and we henceforth assume that (1.2) is a strong integral.

Every researcher in the theory of differential equations and, perhaps more importantly, in applied mathematics, can easily single out many instances of differential equations with important invariants. Familiar examples are

- Conservation of mechanical energy or of volume by a wide range of physical systems;
- Conservation of symplectic form and of Hamiltonian energy by Hamiltonian systems (Sanz-Serna & Calvo 1994);
- Retention of isospectrality by Toda lattice equations (Toda 1981) and by several systems originating in numerical linear algebra (Chu 1994);
- Systems of the form $\boldsymbol{y}' = \boldsymbol{a}(t, \boldsymbol{y})\boldsymbol{y}$, where $\boldsymbol{y}_0$ belongs to a *Lie group* G and $\boldsymbol{a}(t, \boldsymbol{y})$ maps $\mathbb{R}^+ \times G$ to $\mathfrak{g}$, the *Lie algebra* corresponding to G: in that case $\boldsymbol{y}(t) \in G$ for all $t \geq 0$ (Iserles & Nørsett 1997, Zanna 1996);
- Flows on homogeneous spaces, i.e. smooth manifolds which are closed under transitive group action (Munthe-Kaas & Zanna 1997).

We wish to explore in this paper the retention of invariance under discretization. Although the apparent desirability of this practice might be obvious at first glance, it nonetheless calls for justification. The main conclusion of this paper is that many of the more popular discretization methods are utterly useless when it comes to the recovery of invariants. Hence, the desirability of invariance must be often weighed, on a case-by-case basis, against the additional cost of using more 'exotic' time-stepping schemes.

The retention of invariance under discretization is motivated by a number of considerations. Firstly, invariants (also known as conservation laws or integrals) often represent important physical quantities of the underlying system and, intuitively, we should distrust a numerical solution that renders them wrongly. Secondly, retention of invariants is often instrumental in decreasing the accumulation of error in long-term time-stepping discretization (Calvo, Iserles & Zanna 1996, Sanz-Serna & Calvo 1994) and in the recovery of correct asymptotic behaviour (Wisdom & Holman 1991). Thirdly, and perhaps most importantly, it makes sense to be guided by a 'principle of avarice': having deduced, often after prolonged and difficult mathematical *tour de force,* qualitative features of an underlying system, it makes little sense to give them up during discretization ....

Setting $0 = t_0 < t_1 < t_2 < \cdots$, a grid in $\mathbb{R}^+$, we let

$$\boldsymbol{y}_{n+1} = \boldsymbol{\varphi}_{\boldsymbol{f}}(\boldsymbol{y}_{n-m+1}, \boldsymbol{y}_{n-m+2}, \ldots, \boldsymbol{y}_n), \qquad n \geq m-1, \tag{1.3}$$

where $\boldsymbol{y}_l \approx \boldsymbol{y}(t_l)$, $l \in \mathbb{Z}^+$, be an arbitrary m-step discretization method for the ordinary differential system (1.1). Note that (1.3) might originate in an explicit or an implicit scheme, although the latter case may impose restrictions on the grid, to ensure that the underlying algebraic equations possess a unique solution, as represented above.

We say that (1.3) is $\mathcal{M}$-invariant, where $\mathcal{M}$ is a given smooth manifold, if $\boldsymbol{y}(t) \in \mathcal{M}$, $t \geq 0$, and $\boldsymbol{y}_n \in \mathcal{M}$, $n = 0, 1, \ldots, m-1$, imply that $\boldsymbol{y}_n \in \mathcal{M}$ for all $n \in \mathbb{Z}^+$.[1]

The contention of this paper is that many classical numerical methods, not least the ubiquitous *multistep* methods, cannot be $\mathcal{M}$-invariant unless $\mathcal{M}$ is a linear manifold (i.e., the function ρ is linear in the components of $\boldsymbol{y}$). In Section 2 we introduce Taylor-type methods (1.3), accompanying our exposition by a number of examples. Subsequently, in Section 3 we prove that an $\mathcal{M}$-invariance of a Taylor-type method is possible only if the function ρ obeys a specific nonlinear degenerate parabolic differential equation, the *Bateman equation.* The level sets of the Bateman equation are the subject of Section 4, where we demonstrate that they are precisely all linear manifolds in $\mathbb{R}^d$. Therefore, no Taylor-type method may be $\mathcal{M}$-invariant on a nonlinear manifold.

Note that the poor retention of symplecticity, arguably the most interesting example of invariance, by multistep methods has been already noted by Hairer & Leone (1997), who have used an entirely different analytic approach.

Not wishing to conclude our paper on a pessimistic note, we mention in Section 5 a number of methods, some classical and others very recent, that are $\mathcal{M}$-invariant for various manifolds $\mathcal{M}$.

2 Taylor-type time-stepping methods

We say that the m-step discretization method (1.3) of order p is of *Taylor type* if, assuming $\boldsymbol{y}_l = \boldsymbol{y}(t_l)$ for $l = n-m+1, n-m+2, \ldots, n$, $n \geq m-1$, it is true that

$$\boldsymbol{y}_{n+1} - \boldsymbol{y}(t_{n+1}) = c_n h_n^{p+1} \boldsymbol{y}^{(p+1)}(t_n) + \mathcal{O}\big(h_n^{p+2}\big), \tag{2.1}$$

where $h_n = t_{n+1} - t_n$ and the constant $c_n \neq 0$ does not depend on the ordinary differential system (1.1).[2] Equivalently, the method is of Taylor type if for all $n \geq m-1$ it is true that

$$\boldsymbol{y}(t_{n+1}) = \boldsymbol{\varphi}_{\boldsymbol{f}}(\boldsymbol{y}(t_{n-m+1}), \boldsymbol{y}(t_{n-m+2}), \ldots, \boldsymbol{y}(t_n)) + c_n h_n^{p+1} \boldsymbol{y}^{(p+1)}(t_n) + \mathcal{O}\big(h_n^{p+2}\big).$$

The simplest example of a Taylor-type method is the *Taylor method.* Given a smooth function $\boldsymbol{f}$, we can repeatedly differentiate (1.1) to obtain expressions for progressively higher derivatives of $\boldsymbol{y}$,

$$\boldsymbol{y}^{(r)} = \boldsymbol{g}_r(t, \boldsymbol{y}), \qquad r \in \mathbb{Z}^+.$$

Thus,

$$\begin{aligned} \boldsymbol{g}_0(t, \boldsymbol{y}) &= \boldsymbol{y}, \\ \boldsymbol{g}_1(t, \boldsymbol{y}) &= \boldsymbol{f}(t, \boldsymbol{y}), \\ \boldsymbol{g}_2(t, \boldsymbol{y}) &= \frac{\partial \boldsymbol{f}(t, \boldsymbol{y})}{\partial t} + \frac{\boldsymbol{f}(t, \boldsymbol{y})}{\partial \boldsymbol{y}} \boldsymbol{f}(t, \boldsymbol{y}) \end{aligned}$$

[1]This concept should not be confused with the G-invariance of a manifold with regard to the action of a Lie group G (Olver 1995).

[2]This definition is considerably more general than in an earlier paper, (Iserles 1997*b*), where we have presented, employing a different technique, negative results on $\mathcal{M}$-invariance capabilities of *explicit* methods.

and so on. A p-th order Taylor method is based on truncating the Taylor series at t_n,

$$\boldsymbol{y}_{n+1} = \sum_{r=0}^{p} \frac{1}{r!} \boldsymbol{g}_r(t_n, \boldsymbol{y}_n),$$

and it is an immediate consequence of the Taylor theorem that, locally,

$$\boldsymbol{y}_{n+1} - \boldsymbol{y}(t_{n+1}) = \frac{1}{(p+1)!} h_n^{p+1} \boldsymbol{y}^{(p+1)}(t_n) + \mathcal{O}\big(h_n^{p+2}\big) .$$

Another example of a Taylor-type method is a *multistep* scheme

$$\sum_{l=0}^{m} \rho_l \boldsymbol{y}_{n-m+l+1} = h \sum_{l=0}^{m} \sigma_l \boldsymbol{f}(t_{n-m+l+1}, \boldsymbol{y}_{n-m+l+1}), \tag{2.2}$$

where $t_n = nh$, $n \in \mathbb{Z}^+$ and $\rho_m = 1$. We are assuming here that the polynomials

$$\tilde{\rho}(w) = \sum_{l=0}^{m} \rho_l w^l \qquad \text{and} \qquad \tilde{\sigma}(w) = \sum_{l=0}^{m} \sigma_l w^l$$

are relatively prime and that $\tilde{\rho}$ obeys the *root condition* (all its zeros are in $|w| \le 1$ and the zeros on the unit circle are simple). Note that (2.2) is not in the form (1.3), but it can be converted to it by the implicit function theorem, provided that $h > 0$ is sufficiently small.

The method (2.2) is of order p whenever

$$\tilde{\rho}(\mathrm{e}^z) - z\tilde{\sigma}(\mathrm{e}^z) = cz^{p+1} + \mathcal{O}\big(z^{p+1}\big), \qquad z \to 0,$$

where $c \neq 0$. In that case it is trivial to verify that the method obeys (2.1) with $c_n = c$ and is indeed of Taylor type.

Our last example generalises both Taylor and multistep methods to *multistep–multiderivative* methods

$$\sum_{l=0}^{m} \sum_{r=0}^{q} \alpha_{l,r} h^r \boldsymbol{g}_r(t_{n-m+l+1}, \boldsymbol{y}_{n-m+l+1}) = 0.$$

It is left to the reader to verify that this is indeed, subject to the satisfaction of the root condition by the polynomial $\sum_{l=0}^{m} \alpha_{l,0} w^l$, a Taylor-type method.

It is perhaps more insightful to comment on which methods are *not* of Taylor type. The simplest example is presented by *Runge–Kutta* schemes. Thus, for example, applying the two-stage second-order explicit Runge–Kutta scheme

$$\begin{aligned} \boldsymbol{k}_1 &= \boldsymbol{f}(t_n, \boldsymbol{y}_n), \\ \boldsymbol{k}_2 &= \boldsymbol{f}(t_n + \tfrac{2}{3} h_n, \boldsymbol{y}_n + \tfrac{2}{3} h_n \boldsymbol{k}_1) \\ \boldsymbol{y}_{n+1} &= \boldsymbol{y}_n + h_n (\tfrac{1}{4} \boldsymbol{k}_1 + \tfrac{3}{4} \boldsymbol{k}_2) \end{aligned}$$

to the scalar, autonomous equation $y' = f(y)$ results in the error expansion

$$y_{n+1} - y(t_{n+1}) = \frac{h_n^3}{6} [f'(y_n)]^2 f(y_n) + \mathcal{O}\big(h_n^4\big)$$

and this form is inconsistent with (2.1). This example is typical of Runge–Kutta and many other *general linear* methods (Butcher 1987, Hairer & Wanner 1991). Such methods call for significantly more advanced order analysis, e.g. by using B-series (Hairer & Wanner 1991, Sanz-Serna & Calvo 1994), precisely because they are not of Taylor type. However, one of the main conclusions of this paper is that this feature, a frequent source for frustration, represents an underlying virtue of such methods when the retention of invariance is at issue.

3 A necessary condition for $\mathcal{M}$-invariance

Let us suppose that a Taylor-type method (1.3) is $\mathcal{M}$-invariant. Therefore, according to (1.2), $\rho(\boldsymbol{y}_n) = 0$, $n \in \mathbb{Z}^+$, where, for clarity's sake, we suppress the dependence of ρ on $\boldsymbol{y}_0$. According to (2.1), we thus have

$$\begin{aligned} 0 = \rho(\boldsymbol{y}_{n+1}) &= \rho(\boldsymbol{y}(t_{n+1}) + ch_n^{p+1}\boldsymbol{y}^{(p+1)}(t_n) + \mathcal{O}(h_n^{p+2})) \\ &= \rho(\boldsymbol{y}(t_n)) + ch_n^{p+1} \sum_{l=1}^{d} \frac{\partial \rho(\boldsymbol{y}(t_n))}{\partial y_l} y_l^{(p+1)}(t_n) + \mathcal{O}(h_n^{p+2}) \,. \end{aligned}$$

Recall, however, that the exact solution of (1.1) resides in $\mathcal{M}$, thus, letting $h_n \to 0$, we deduce that a necessary condition for $\mathcal{M}$-invariance is

$$\sum_{l=1}^{d} \frac{\partial \rho(\boldsymbol{y}(t))}{\partial y_l} y_l^{(p+1)}(t) \equiv 0. \tag{3.1}$$

Note that the invariance of (1.1) implies that (3.1) is satisfied with $p = 0$. Our conclusion is that the retention of invariance under discretization requires the satisfaction of (3.1) for a positive value of p. Interestingly enough, (3.1) has nothing to do with the specific form of the method (1.3), except for its order. We devote the remainder of this section, as well as Section 4, to the exploration of manifolds $\mathcal{M}$ such that every flow $\boldsymbol{y}$ on $\mathcal{M}$ obeys the condition (3.1).

Any system (1.1) which is invariant on $\mathcal{M}$ can be written in a *skew-gradient* form,

$$\boldsymbol{y}' = S(t, \boldsymbol{y})\nabla\rho(\boldsymbol{y}) \tag{3.2}$$

where $S(\,\cdot\,,\,\cdot\,)$ is a $d \times d$ skew-symmetric matrix (Itoh & Abe 1988, Quispel & Capel 1997). As a matter of fact, the form (3.2) characterises all ordinary differential systems which are invariant on the manifold (1.2).

We henceforth restrict our attention to $d = 2$. This, as will be made clear in the sequel, does not lead to any loss of generality, whilst simplifying the exposition a great deal. Given two dimensions, (3.2) becomes

$$\begin{aligned} y_1' &= \psi(t, \boldsymbol{y}) \frac{\partial \rho(\boldsymbol{y})}{\partial y_2}, \\ y_2' &= -\psi(t, \boldsymbol{y}) \frac{\partial \rho(\boldsymbol{y})}{\partial y_1}, \end{aligned} \tag{3.3}$$

where the smooth function $\psi : \mathbb{R}\times\mathbb{R}^2 \to \mathbb{R}$ is arbitrary. (The case $\psi \equiv 1$ corresponds to Hamiltonian systems, whereby ρ is the Hamiltonian energy, but this observation plays no part in our analysis.) Note that $\mathcal{M}$-invariance means that a numerical method stays on $\mathcal{M}$ for *all* choices of ψ. In other words, in order to demonstrate that a method is not $\mathcal{M}$-invariant, it is enough to prove that there exists a single such function ψ for which the identity (3.1), a necessary condition for $\mathcal{M}$-invariance, is not satisfied.

Proposition 1 *Let $\psi \not\equiv 0$ depend only on t. Then, for every $l \in \mathbb{N}$, there exist functions $\xi_1^{[l]}$ and $\xi_2^{[l]}$, dependent only on $\psi, \psi', \ldots, \psi^{(l-3)}$, such that*

$$\begin{aligned} y_1^{(l)} &= \psi^{(l-1)}\frac{\partial\rho}{\partial y_2} + (l-1)\psi\psi^{(l-2)}\left(\frac{\partial^2\rho}{\partial y_1\partial y_2}\frac{\partial\rho}{\partial y_2} - \frac{\partial^2\rho}{\partial y_2^2}\frac{\partial\rho}{\partial y_1}\right) + \xi_1^{[l]}, \\ y_2^{(l)} &= -\psi^{(l-1)}\frac{\partial\rho}{\partial y_1} + (l-1)\psi\psi^{(l-2)}\left(\frac{\partial^2\rho}{\partial y_1\partial y_2}\frac{\partial\rho}{\partial y_1} - \frac{\partial^2\rho}{\partial y_1^2}\frac{\partial\rho}{\partial y_2}\right) + \xi_2^{[l]}. \end{aligned} \tag{3.4}$$

Proof The formula (3.4) is certainly true (with $\xi_1^{[1]}, \xi_2^{[1]} \equiv 0$) for $l = 1$, since it coincides with (3.3). Likewise, direct differentiation affirms that

$$\begin{aligned} y_1'' &= \psi'\frac{\partial\rho}{\partial y_2} + \psi\left(\frac{\partial^2\rho}{\partial y_2\partial y_1}y_1' + \frac{\partial^2\rho}{\partial y_2^2}y_2'\right) = \psi'\frac{\partial\rho}{\partial y_2} + \psi^2\left(\frac{\partial^2\rho}{\partial y_1\partial y_2}\frac{\partial\rho}{\partial y_2} - \frac{\partial^2\rho}{\partial y_2^2}\frac{\partial\rho}{\partial y_1}\right) \\ y_2'' &= -\psi'\frac{\partial\rho}{\partial y_1} - \psi\left(\frac{\partial^2\rho}{\partial y_1^2}y_1' + \frac{\partial^2\rho}{\partial y_1\partial y_2}\right) = -\psi'\frac{\partial\rho}{\partial y_1} + \psi^2\left(\frac{\partial^2\rho}{\partial y_1\partial y_2}\frac{\partial\rho}{\partial y_1} - \frac{\partial^2\rho}{\partial y_1^2}\frac{\partial\rho}{\partial y_2}\right). \end{aligned}$$

Thus, $\xi_1^{[2]}, \xi_2^{[2]} \equiv 0$, consistently with the statement of the proposition.

We continue by induction. Differentiating the first equation in (3.4), we obtain

$$y_1^{(l+1)} = \psi^{(l)}\frac{\partial\rho}{\partial y_2} + l\psi\psi^{(l-1)}\left(\frac{\partial^2\rho}{\partial y_1\partial y_2}\frac{\partial\rho}{\partial y_2} - \frac{\partial^2\rho}{\partial y_2^2}\frac{\partial\rho}{\partial y_1}\right) + \xi_1^{[l+1]},$$

where

$$\begin{aligned} \xi_1^{[l+1]} &= \xi_1^{[l]\prime} + (l-1)\psi'\psi^{(l-2)}\left(\frac{\partial^2\rho}{\partial y_1\partial y_2}\frac{\partial\rho}{\partial y_2} - \frac{\partial^2\rho}{\partial y_2^2}\frac{\partial\rho}{\partial y_1}\right) \\ &\quad + (l-1)\psi^2\psi^{(l-2)}\left\{\frac{\partial^3\rho}{\partial y_1^2\partial y_2}\left(\frac{\partial\rho}{\partial y_2}\right)^2 - 2\frac{\partial^3\rho}{\partial y_1\partial y_2^2}\frac{\partial\rho}{\partial y_1}\frac{\partial\rho}{\partial y_2} + \frac{\partial^3\rho}{\partial y_2^3}\left(\frac{\partial\rho}{\partial y_1}\right)^2\right. \\ &\quad \left. + \left[\left(\frac{\partial^2\rho}{\partial y_1\partial y_2}\right)^2 - \left(\frac{\partial^2\rho}{\partial y_1^2}\right)^2\right]\frac{\partial\rho}{\partial y_2} - \frac{\partial^2\rho}{\partial y_1\partial y_2}\left(\frac{\partial^2\rho}{\partial y_2^2} - \frac{\partial^2\rho}{\partial y_1^2}\right)\frac{\partial\rho}{\partial y_1}\right\}. \end{aligned}$$

The complexity of the last expression notwithstanding, it is clear from the induction hypothesis that $\xi_1^{[l+1]}$ depends just on $\psi, \psi', \ldots, \psi^{(l-2)}$ and is independent of $\psi^{(i)}$ for $i \geq l-1$. An identical argument extends to $\xi_2^{[l+1]}$ and the proposition is true. □

We deduce from (3.4) that

$$y_1^{(p+1)}\frac{\partial\rho}{\partial y_1} + y_2^{(p+1)}\frac{\partial\rho}{\partial y_2} = p\psi\psi^{(p-1)}\mathcal{B}(\rho) + \xi^{[p+1]}, \tag{3.5}$$

where

$$\mathcal{B}(u) = \left(\frac{\partial u}{\partial y_2}\right)^2 \frac{\partial^2 u}{\partial y_1^2} - 2\frac{\partial u}{\partial y_1}\frac{\partial u}{\partial y_2}\frac{\partial^2 u}{\partial y_1 \partial y_2} + \left(\frac{\partial u}{\partial y_1}\right)^2 \frac{\partial^2 u}{\partial y_2^2}$$

and the function $\xi^{[p+1]}$ is independent of $\psi^{(i)}$ for $i \geq p-1$.

Let us assume that the function ρ is not a solution of the *Bateman* equation

$$\mathcal{B}(\rho) = \left(\frac{\partial \rho}{\partial y_2}\right)^2 \frac{\partial^2 \rho}{\partial y_1^2} - 2\frac{\partial \rho}{\partial y_1}\frac{\partial \rho}{\partial y_2}\frac{\partial^2 \rho}{\partial y_1 \partial y_2} + \left(\frac{\partial \rho}{\partial y_1}\right)^2 \frac{\partial^2 \rho}{\partial y_2^2} = 0. \tag{3.6}$$

in some nonempty neighbourhood $\mathcal{V}$ of $\mathbb{R}^2$. Therefore there exists an open, nonempty set $\mathcal{W} \subseteq \mathcal{V}$ where $\mathcal{B}(\rho)$ is bounded away from zero.

Note that (3.6) is sometimes known in the literature as the *Born–Infeld* equation (Whitham 1974), but we adhere in this paper to the terminology of Fairlie & Mulvey (1994), where the sobriquet 'Born–Infeld equation' is reserved for a more general construct.

Recall that we have still retained a complete freedom to choose the function ψ, subject to its smoothness. We exercise this freedom by letting ψ be a polynomial of degree $(p-1)$, of the form

$$\psi(t) = 1 + ct^{p-1},$$

where c is an arbitrary real constant such that

$$c > \frac{p}{(p+1)!} \times \frac{|\xi^{[p+1]}(t)|}{\min_{\boldsymbol{y}\in\mathcal{W}} |\mathcal{B}(\rho(\boldsymbol{y}))|}$$

for $t \in [0, t^*]$.[3] Substituting in (3.5), we deduce that

$$\begin{aligned} &\left| y_1^{(p+1)} \frac{\partial \rho}{\partial y_1} + y_2^{(p+1)} \frac{\partial \rho}{\partial y_2} \right| \\ &> (p-1)!(p+1)c\psi(t)|\mathcal{B}(\rho(\boldsymbol{y}))| - |\xi^{[p+1]}(t)| > 0, \qquad t \in [0,t^*], \quad \boldsymbol{y} \in \mathcal{W}. \end{aligned} \tag{3.7}$$

Lemma 2 *Let $\mathcal{M}$ be a smooth manifold in $\mathbb{R}^d$ and suppose that there exists a two-dimensional slice $\mathcal{N}$ through $\mathcal{M}$ (i.e., $\mathcal{N}$ is an intersection of $\mathcal{M}$ with a two-dimensional linear subspace of $\mathbb{R}^d$ which, without loss of generality, we may assume to be $(y_1, y_2, 0, \ldots, 0) \in \mathbb{R}^d$) such that $\mathcal{N} = \{\boldsymbol{x} \in \mathbb{R}^2 : \rho(\boldsymbol{x}) = 0\}$ and the function ρ does not obey the Bateman equation (3.6). In that case no Taylor-type method can be $\mathcal{M}$-invariant.*

Proof It is enough to prove that no Taylor-type method is $\mathcal{N}$-invariant, since $\mathcal{N}$ is a smooth submanifold of $\mathcal{M}$. For suppose that a method is not $\mathcal{N}$-invariant, hence the solution sequence $\{\boldsymbol{y}_n\}$ departs from $\mathcal{N}$ for some system (1.1) whose exact solution obeys $\rho(\boldsymbol{y}(t)) \equiv 0$. By design, the solution always stays in the two-dimensional linear space spanned by the first two coordinates in $\mathbb{R}^d$. Therefore, either $\mathcal{M}$ is in the hyperplane – and this contradicts our construction – or $\boldsymbol{y}_n \notin \mathcal{M}$ for some $n \in \mathbb{Z}^+$.

[3]Hence, (3.3) becomes a $(p-1)$st-order perturbation of a Hamiltonian system

The proof now follows readily from Proposition 1 and the inequality (3.7). □

We deduce that conservation of invariance by Taylor-type methods is closely related to the geometry of the underlying manifold $\mathcal{M}$ and that the link is furnished by the Bateman equation.

4 The Bateman equation and invariance on nonlinear manifolds

The main result of the last section, Lemma 2, is that $\mathcal{M}$-invariance of a Taylor method is possible only if every two-dimensional slice through $\mathcal{M}$ is a level set of a solution of the Bateman equation (3.6). In the present section we explore level sets of (3.6), proving that they are of an excessively simple form.

Let us suppose that the analytic function $\rho \not\equiv 0$ is a solution of (3.6) for all $\boldsymbol{y} \in \mathcal{W}$, where $\mathcal{W}$ is a nonempty, open neighbourhood in $\mathbb{R}^2$. There are two possibilities: either $\partial\rho(\boldsymbol{y})/\partial y_2 \equiv 0$ for all $\boldsymbol{y} \in \mathcal{W}$ or there exists an open, nonempty set $\mathcal{X} \subseteq \mathcal{W}$ such that $\partial\rho(\boldsymbol{y})/\partial y_2 \neq 0$ for all $\boldsymbol{y} \in \mathcal{X}$. In the first case, which is consistent with ρ being a solution of (3.6), analyticity of ρ implies that, except for the case when $\mathcal{M}$ consists of a single point (hence, is not a manifold!), it must be a linear function, $\rho(\boldsymbol{y}) = \omega(\alpha + \beta y_1)$, say, where ω is an arbitrary analytic function (which makes no difference to the geometry of level sets). In the second case we use the implicit function theorem to deduce that there exists a smooth function $\eta : \mathbb{R} \to \mathbb{R}$ such that

$$\rho(\boldsymbol{y}) = 0 \qquad \Rightarrow \qquad y_2 = \eta(y_1).$$

We differentiate $\rho(x, \eta(x)) = 0$, the outcome being

$$\frac{\partial\rho(x,\eta(x))}{\partial y_1} + \frac{\partial\rho(x,\eta(x))}{\partial y_2}\eta'(x) = 0.$$

This yields an explicit expression for the derivative of η,

$$\eta'(x) = -\frac{\partial\rho(x,\eta(x))}{\partial y_1} \bigg/ \frac{\partial\rho(x,\eta(x))}{\partial y_2}. \tag{4.1}$$

Differentiating again, we have

$$\frac{\partial^2\rho(x,\eta(x))}{\partial y_1^2} + 2\frac{\partial^2\rho(x,\eta(x))}{\partial y_1 \partial y_2}\eta'(x) + \frac{\partial^2\rho(x,\eta(x))}{\partial y_2^2}[\eta'(x)]^2 + \frac{\partial\rho(x,\eta(x))}{\partial y_2}\eta''(x) = 0.$$

Substitution of (4.1) results in

$$\mathcal{B}(\rho(x,\eta(x))) + \left(\frac{\partial\rho(x,\eta(x))}{\partial y_2}\right)^3 \eta''(x) = 0$$

and, ρ being a solution of (3.6), $(x, \eta(x)) \in \mathcal{X}$ implies $\eta''(x) = 0$.

The satisfaction of $\eta'' \equiv 0$ in $\mathcal{X}$ means that η is a linear function, therefore $\rho(\boldsymbol{y}) = \omega(\alpha + \beta y_1 + \gamma y_2)$, with a linear level set. We thus deduce that also in this case the level sets of ρ are necessarily linear.

Theorem 3 *A Taylor-type method is $\mathcal{M}$-invariant if and only if $\mathcal{M}$ is a linear manifold.*

Proof Suppose that $\mathcal{M}$ is nonlinear. In that case there exists a two-dimensional slice $\mathcal{N}$ through $\mathcal{M}$ which is also nonlinear. According to our analysis, $\mathcal{N}$ cannot be a level set of a solution of the Bateman equation, therefore, as a consequence of Lemma 2, no Taylor-type method can be $\mathcal{M}$-invariant.

To complete the proof, we observe that every consistent time-stepping method, and in particular every Taylor-type method, is $\mathcal{M}$-invariant on every linear manifold $\mathcal{M}$. □

5 Conclusions

The practical consequence of our main result, Theorem 3, is clear. Taylor-type methods, and in particular multistep methods, are worthless, insofar as correct rendition of nonlinear invariants is concerned. This is not to deny the many important uses of multistep methods in the numerical analysis of ordinary differential equations: as we have already mentioned in Section 1, retention of invariance is just one consideration among many. Yet, whenever it makes good sense to respect invariants, we need to resort to different methods.

As we have already mentioned in Section 2, Runge–Kutta methods are not of Taylor type, hence they are not a subject to the restrictions of Theorem 3. Indeed, Cooper (1987) proved that any ν-stage Runge–Kutta method

$$\begin{aligned} \boldsymbol{\phi}_l &= \boldsymbol{y}_n + h\sum_{j=1}^{\nu} a_{l,j}\boldsymbol{k}_j, \\ \boldsymbol{k}_l &= \boldsymbol{f}(t_n + c_l h, \boldsymbol{\phi}_l), \end{aligned} \qquad l = 1, 2, \ldots, \nu,$$

$$\boldsymbol{y}_{n+1} = \boldsymbol{y}_n + h\sum_{l=1}^{\nu} b_l\boldsymbol{k}_l$$

is invariant on *every* quadratic manifold

$$\mathcal{M} = \{\boldsymbol{y} \in \mathbb{R}^d \,:\, \boldsymbol{y}^{\mathrm{T}} P\boldsymbol{y} + \boldsymbol{q}^{\mathrm{T}}\boldsymbol{y} + r = 0\}$$

if

$$b_l a_{l,j} + b_j a_{j,l} = b_l b_j, \qquad l, j = 1, 2, \ldots, \nu. \tag{5.1}$$

It has been proved in (Calvo, Iserles & Zanna 1997*a*) that the condition (5.1) is also necessary for invariance on quadratic manifolds. A forthcoming paper will address itself to invariance of Runge–Kutta methods and its main conclusion is that not much, besides quadratic invariance, can be expected from such schemes.

Quadratic manifolds are of great importance in a wide range of applications, e.g. orthogonal flows (Dieci, Russell & van Vleck 1994), Hamiltonian problems (Sanz-Serna & Calvo 1994), equations on the orthogonal groups O(d) and SO(d), on the unitary group $U(d)$ and on Stiefel and Grassmann manifolds (Olver 1995).[4] Moreover, certain invariants can be represented as a group action of $SO(d)$, and this can be exploited in their discretization by Runge–Kutta methods, isospectral flows being the most prominent example (Calvo et al. 1996, Calvo et al. 1997*a*, Calvo, Iserles & Zanna 1997*b*).

The last few years have seen the emergence of a new breed of numerical schemes, which do not fit conveniently into the classical framework of general linear methods, and which have been developed with the explicit purpose of rendering correctly invariant manifolds. It is not the purpose of this paper to review such methods and, anyway, it is probably premature to attempt such an endeavour. Having said this, we briefly mention a number of novel methods:

1. The most obvious approach to the maintenance of an invariant is to reformulate the underlying system as a *differential-algebraic equation* (Ascher 1997). This approach holds a number of attractions but it is hardly a remedy for all situations (cf. a discussion in (Iserles 1997*a*)).

2. Itoh & Abe (1988) have introduced a first-order weakly implicit method, based on the skew-gradient representation (3.2), which can be made invariant whenever ρ is known. This method has been extended by Quispel & Capel (1997).

3. Crouch & Grossman (1993) introduced methods which integrate along rigid frames in the tangent bundle of the underlying manifold. Such methods are restricted by the need to know the tangent bundle explicitly (an easy task in the case of Lie groups) and they lead to complicated order conditions (Owren & Marthinsen 1997).

4. An ingenious extension of Runge–Kutta methods to Lie groups has been introduced by Munthe-Kaas (1995) and further refined in (Munthe-Kaas 1997).

5. Several novel methods solve differential systems on Lie groups by expanding in the underlying Lie algebra and translating back to the Lie group by means of the exponential map (Iserles & Nørsett 1997, Zanna 1996).

6. This brief list will be incomplete without a recent technique which, alternating between an original and an 'adjoint' equation, progressively decreases the discrepancy of an arbitrary numerical method from an invariant manifold. This idea has been originally implemented for orthogonal and isospectral flows (Calvo et al. 1997*b*, Iserles 1997*a*) and has been extended to arbitrary involutory automorphisms and anti-automorphisms (Iserles, McLachlan & Zanna 1997).

It is premature to debate the relative advantages and disadvantages of the above approaches.

[4] In the case of Hamiltonian equations, the symplectic manifold is quadratic – such systems possess further invariants, e.g. Hamiltonian energy, which in general are not quadratic.

References

Ascher, U. M. (1997), 'Stabilization of invariants of discretized differential systems', *Num. Algorithms*. to appear.

Butcher, J. C. (1987), *The Numerical Analysis of Ordinary Differential Equations*, John Wiley, Chichester.

Calvo, M. P., Iserles, A. & Zanna, A. (1996), 'Runge–Kutta methods for orthogonal and isospectral flows', *Appl. Num. Maths* **22**, 153–163.

Calvo, M. P., Iserles, A. & Zanna, A. (1997*a*), 'Numerical solution of isospectral flows', *Maths Comput.*. to appear.

Calvo, M. P., Iserles, A. & Zanna, A. (1997*b*), Semi-explicit methods for isospectral flows, Technical report, DAMTP, Univ. Cambridge. to appear.

Chu, M. T. (1994), A list of matrix flows with applications, *in* A. Bloch, ed., 'Hamiltonian and Gradient Flows', Fields Institute Communications, American Maths Soc., Providence, RI, pp. 87–97.

Cooper, G. J. (1987), 'Stability of Runge–Kutta methods for trajectory problems', *IMA J. Num. Anal.* **7**, 1–13.

Crouch, P. & Grossman, R. (1993), 'Numerical integration of ordinary differential equations on manifolds', *J. Nonlinear Sc.* **3**, 1–33.

Dieci, L., Russell, R. D. & van Vleck, E. S. (1994), 'Unitary integrators and applications to continuous orthonormalization techniques', *SIAM J. Numer. Anal.* **31**, 261–281.

Fairlie, D. B. & Mulvey, J. A. (1994), 'Integrable generalisations of the 2-dimensional Born Infeld equation', *J. Phys. A* **27**, 1317–1324.

Guillemin, V. & Pollack, A. (1974), *Differential Topology*, Prentice-Hall, Inc., Englewood Cliffs, NJ.

Hairer, E. & Leone, P. (1997), Order barriers for symplectic multi-value methods, Technical report, Université de Genéve.

Hairer, E. & Wanner, G. (1991), *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems*, Springer-Verlag, Berlin.

Iserles, A. (1997*a*), Beyond the classical theory of computational ordinary differential equations, *in* I. Duff & G. Watson, eds, 'State of the Art in Numerical Analysis', Oxford University Press, Oxford, pp. 171–192.

Iserles, A. (1997*b*), Numerical methods on (and off) manifolds, *in* F. Cucker & M. Shub, eds, 'Foundations of Computational Mathematics', Springer-Verlag, Berlin, pp. 180–189.

Iserles, A. & Nørsett, S. P. (1997), On the solution of linear differential equations in Lie groups, Technical Report NA1997/03, DAMTP, Univ. Cambridge.

Iserles, A., McLachlan, R. I. & Zanna, A. (1997), Approximately preserving symmetries in numerical integration, Technical report, DAMTP, Univ. Cambridge. to appear.

Itoh, T. & Abe, K. (1988), 'Hamiltonian-conserving discrete canonical equations based on variational difference quotients', *J. Comput. Phys.* **76**, 85–102.

Munthe-Kaas, H. (1995), 'Lie–Butcher theory for Runge–Kutta methods', *BIT* **35**, 572–587.

Munthe-Kaas, H. (1997), Runge–Kutta methods on Lie groups, Technical Report 122, Univ. of Bergen.

Munthe-Kaas, H. & Zanna, A. (1997), Numerical integration of differential equations on homogeneous manifolds, *in* F. Cucker & M. Shub, eds, 'Foundations of Computational Mathematics', Springer-Verlag, Berlin, pp. 305–315.

Olver, P. (1995), *Equivalence, Invariants and Symmetry*, Cambridge University Press, Cambridge.

Owren, B. & Marthinsen, A. (1997), Order conditions for integration methods based on rigid frames, Technical Report 1/1997, NTNU, Trondheim.

Quispel, G. R. W. & Capel, H. W. (1997), Solving ODE's numerically while preserving all first integrals, Technical report, La Trobe University.

Sanz-Serna, J. M. & Calvo, M. P. (1994), *Numerical Hamiltonian Problems*, Chapman and Hall, London.

Toda, M. (1981), *Theory of Nonlinear Lattices*, Springer-Verlag, Berlin.

Whitham, G. B. (1974), *Linear and Nonlinear Waves*, John Wiley, New York.

Wisdom, J. & Holman, M. (1991), 'Symplectic maps for the N-body problem', *Astron. J.* **102**, 1528–1538.

Zanna, A. (1996), The method of iterated commutators for ordinary differential equations on Lie groups, Technical Report 1996/NA12, DAMTP, Univ. Cambridge.