Journal of Geometric Mechanics

https://www.aimspress.com/journal/jgm

Journal of Geometric Mechanics, 15(1): 98–115
DOI: 10.3934/jgm.2023005
Received: 12 January 2022
Revised: 12 August 2022
Accepted: 10 September 2022
Published: 08 December 2022

Research article

Backward error analysis for conjugate symplectic methods

Robert I McLachlan[1], **and Christian Offen**[2,*]

[1] Massey University, Private Bag 11 222, Palmerston North 4442, New Zealand
[2] Paderborn University, Warburger Str. 100, 33098 Paderborn, Germany

* **Correspondence:** Email: christian.offen@uni-paderborn.de

Abstract: The numerical solution of an ordinary differential equation can be interpreted as the exact solution of a nearby modified equation. Investigating the behaviour of numerical solutions by analysing the modified equation is known as backward error analysis. If the original and modified equation share structural properties, then the exact and approximate solution share geometric features such as the existence of conserved quantities. Conjugate symplectic methods preserve a modified symplectic form and a modified Hamiltonian when applied to a Hamiltonian system. We show how a blended version of variational and symplectic techniques can be used to compute modified symplectic and Hamiltonian structures. In contrast to other approaches, our backward error analysis method does not rely on an ansatz but computes the structures systematically, provided that a variational formulation of the method is known. The technique is illustrated on the example of symmetric linear multistep methods with matrix coefficients.

Keywords: variational integrators, backward error analysis, Euler–Lagrange equations, multistep methods, conjugate symplectic methods
Mathematics Subject Classification: Primary: 65D30, 70H25; Secondary: 70H50

1. Introduction

While the forward error of a numerical method compares the exact solution of an ODE with the numerical solution after one time-step h, to obtain qualitative statements about the long-term behaviour of numerical solutions to ODEs, it is helpful to consider a *modified ODE* whose exact solution closely approximates the numerical flow map at grid points. A modified equation can be obtained as an expansion of the numerical solution as a power series in the step-size h. Though the series does not converge in general, optimal truncation techniques have been established such that the flow of the modified system and the numerical method coincide at grid points up to exponentially small error terms.

Computing and analysing the structural properties of modified equations is known as *backward error analysis (BEA)* (see, for instance, [3, §IX] or [6, §5]). Next to the analysis of long-term behaviour of numerical schemes, backward error analysis has been used to improve the initialisation of multi-step methods [2] as well as to improve physics informed machine learning techniques [13, 12, 10].

If a Hamiltonian ODE is discretised by a symplectic integrator, then any truncation of the modified equation is itself a Hamiltonian system with respect to the original symplectic structure and a modified Hamiltonian. These are also called *Shadow Hamiltonians*. The existence of a modified Hamiltonian or a modified Lagrangian is a key ingredient to obtain statements about long-term behaviour of symplectic method, such as oscillatory energy errors over exponentially long time intervals. Moreover, just as in the exact system, symplectic symmetries of the modified system yield conserved quantities for the modified dynamics by Noether's theorem. This explains why symplectic integrators behave well on completely integrable systems. A detailed discussion can be found in [3]. Vermeeren observed that backward error analysis for variational integrators can be done entirely on the Lagrangian side [15].

In contrast to symplectic methods, conjugate symplectic methods preserve a modified symplectic structure rather than the original symplectic structure. Conjugate symplectic methods share the excellent long-term behaviour of symplectic methods. Moreover, Noether's theorem applies such that symmetries of the modified system yield modified conserved quantities of the modified dynamics. When modified structures are explicitly known, explicit expressions of modified conserved quantities can be derived. This motivates the development of techniques to compute modified symplectic structures and Hamiltonians.

While traditional methods for the computation of modified Hamiltonians use an Ansatz (i.e. an educated guess) of the Hamiltonian as a power series and match terms, working with an Ansatz is challenging when a modified Hamiltonian *and* a modified symplectic structure need to be computed simultaneously: the components of matrices representing symplectic structures are in this context not constant but depend on the state space variables, fulfil a symmetry condition, and satisfy the Jacobi identity, which is a partial differential equation [3, §VII.2]. This makes finding a suitable Ansatz difficult.

A typical strategy [3, 8] to obtain a structure preserving numerical method is to approximate the variational principle

$$\delta S = 0, \quad S(y) = \int_{t_0}^{t_N} L(y(t), \dot{y}(t))\mathrm{d}t, \quad y(t_0) = y_0, y(t_N) = y_N \tag{1.1}$$

which governs the exact Euler–Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{y}} - \frac{\partial L}{\partial y} = 0$$

by a discrete variational principle

$$\nabla S_\Delta(\{y_i\}_i) = 0, \quad S_\Delta(\{y_i\}_i) = \sum_{i=0}^{N-1} hL_\Delta(y_i, y_{i+1}).$$

Since y_0 and y_N are fixed in the variations considered in (1.1), the gradient above is taken with respect to all inner grid points $y_1, \ldots, y_{N-1}$. We obtain the discrete Euler–Lagrange equations

$$\mathrm{D}_1 L_\Delta(y_i, y_{i+1}) + \mathrm{D}_2 L_\Delta(y_{i-1}, y_i) = 0, \ i = 0, \ldots, N-1$$

which yield approximations $y_i \approx y(t_0 + ih)$ to an exact solution y. The term recursion is called a *variational method*. Indeed, the class of variational methods is equivalent to the class of symplectic integrators [3, 8].

While backward error analysis for discrete Lagrangians $L_\Delta(y_i, y_{i+1})$ are established, discrete Lagrangians depending on several grid-points $L_\Delta(y_i, y_{i+1}, \dots, y_{i+s})$ corresponding to multistep methods require different approaches because they are not symplectic but only preserve a modified symplectic structure. In other words, these methods are conjugate to symplectic methods. However, the modified symplectic structures or conjugacies, respectively, are given by a formal power series that might not be convergent. Although rigorous optimal truncation results are not available, we will demonstrate in numerical examples that truncations can be useful objects in the analysis of the numerical methods.

In the following, we will introduce *blended backward error analysis* to systematically compute modified Hamiltonian and modified symplectic structures. We will prove the following theorem which applies, for instance, to series expansions $\mathcal{L}_\Delta$ of consistent discrete Lagrangians $L_\Delta(y(t), y(t+h), \dots, y(t+sh))$.

Theorem 1.1. *Consider a power series $\mathcal{L}_\Delta(y^{[\infty]})$ in a formal variable h. The series depends on the jet $y^{[\infty]} = (y, \dot{y}, \ddot{y}, \dots)$ of a variable y such that any truncation only depends on a finite jet of y. Assume further that the truncation to zeroth order constitutes a regular Lagrangian $L(y, \dot{y})$, i.e. $\left(\frac{\partial^2 L}{\partial \dot{y}^i \partial \dot{y}^j}\right)_{i,j}$ is invertible.*

- *There exists a 2nd order modified equation given as a formal power series in h such that for any $N \in \mathbb{N}$ a solution of the modified equation truncated to order $O(h^N)$ solves the Euler–Lagrange equations to $\mathcal{L}_\Delta^{[N]}$ up to an error of order $O(h^{N+1})$, where $\mathcal{L}_\Delta^{[N]}$ is the truncation of $\mathcal{L}_\Delta$ to order $O(h^N)$.*
- *If we denote $z = (y, \dot{y})$, there exists a symplectic structure matrix J_{mod} and a Hamiltonian H_{mod} given as formal power series in h such that solutions to Hamilton's equations*

$$\dot{z} = J_{\mathrm{mod}}^{[N]}(z)^{-1} \nabla H_{\mathrm{mod}}^{[N]}(z)$$

fulfil the modified equation. Here, $J_{\mathrm{mod}}^{[N]}$ and $H_{\mathrm{mod}}^{[N]}$ are truncations to order $O(h^N)$ of J_{mod} and H_{mod}, respectively such that $\mathcal{L}_\Delta^{[N]}$ is regular, i.e. $\left(\frac{\partial^2 \mathcal{L}_\Delta^{[N]}}{\partial y^{(M)i} \partial y^{(M)j}}\right)_{i,j}$ is invertible, where $y^{(M)}$ is the highest derivative of $\mathcal{L}_\Delta^{[N]}$.

The technique will be illustrated for linear multistep methods with matrix valued coefficients (Section 2). These occur, for instance, when in a system of coupled ODEs components of the differential equation are discretised separately with traditional linear multistep method, when multistep methods are stabilised [4, 5, 14], or when discretisation schemes of PDEs are analysed [9].

Moreover, we will analyse under which conditions modified Lagrangians exist: if the original equation is the Euler–Lagrange equation to a variational principle of the form (1.1) for a Lagrangian $L(y, \dot{y})$, it is natural to ask, whether there exists a modified Lagrangian $L_{\mathrm{mod}}(y, \dot{y})$ such that the modified equation is governed by the Euler–Lagrange equations to L_{mod}. In contrast to classical variational integrators, for which $L_{\mathrm{mod}}(y, \dot{y})$ is known to exist and can be computed [15], for conjugate symplectic schemes L_{mod} may only exists in modified variables $L_{\mathrm{mod}}(\tilde{y}, \dot{\tilde{y}})$. We will prove that L_{mod} exists in the original variables $(y, \dot{y})$ if J_{mod} has the form

$$J_{\mathrm{mod}} = \begin{pmatrix} * & * \\ * & 0 \end{pmatrix}.$$

In particular, this shows that for classical consistent, symmetric, linear multistep methods with matrix coefficients with a central force evaluation applied to second order equations L_{mod} exists in the original variables $(y, \dot{y})$ if all coefficients are multiples of the identity matrix, i.e. we have a multistep method with scalar coefficients. However, in the general case of matrix coefficients L_{mod} only exists in modified variables $(\tilde{y}, \dot{\tilde{y}})$.

The article is structured as follows: Section 2 illustrates the ideas of blended backward error analysis on linear multi-step methods. Section 3 shows how to compute the modified data J_{mod} and H_{mod} introduced in Theorem 1.1. The technique is then applied to linear multi-step methods with matrix valued coefficients in Section 4 and results are illustrated by numerical experiments. Additionally, for comparison of blended backward error analysis with classical backward error analysis, **??** contains an application of blended backward error analysis to a mechanical ordinary differential equation discretised with the Störmer-Verlet scheme. A formal proof of Theorem 1.1 is provided in Section 5. Finally, Section 6 discusses the existence of modified Lagrangians as formal power series and future research directions are suggested in Section 7.

2. Application of blended backward error analysis to linear multistep methods with matrix coefficients

To illustrate the idea of blended backward error analysis, we compute modified symplectic structures and Hamiltonians of linear multistep methods.

Consistent, symmetric linear multistep methods with a single force evaluation applied to the second order ODE

$$\ddot{y} = \nabla U(y(t)) \tag{2.1}$$

take the form

$$\sum_{j=1}^{\frac{s}{2}} A_j (y(t - jh) - 2y(t) + y(t + jh)) = h^2 \nabla U(y(t)) \tag{2.2}$$

with

$$\sum_{j=1}^{\frac{s}{2}} j^2 A_j = I. \tag{2.3}$$

Relation (2.3) is coming from the consistency requirement [4]. These are s-step methods, where s is even. Here we allow matrix valued coefficients A_j [4, 5, 9, 14], h is the step-size and I denotes the identity matrix. If the coefficients A_j are scalars, then the schemes constitute classical consistent, symmetric linear multistep methods. A series expansion of (2.2) in h is equivalent to a power series expression of the form

$$\ddot{y} = \nabla U(y(t)) + \sum_{i=1}^{N} h^i \tilde{g}_i(y(t), \dots, y^{(a_i)}) + O(h^{N+1}) \tag{2.4}$$

for some h-independent expressions $\tilde{g}_i$ in the a_i-jet of y at t. Substituting 3rd and higher derivatives on the right hand side of (2.4) with derivatives of (2.4) itself iteratively yields an equation of the form

$$\ddot{y} = \nabla U(y(t)) + \sum_{i=1}^{N} h^i g_i(y(t), \dot{y}(t)) + O(h^{N+1}) \tag{2.5}$$

which is called the *modified equation* of method (2.2) applied to (2.1). We refer to [3] for optimal truncation techniques and a discussion of spurious solutions not covered by the considered modified system for the case of linear multistep methods with scalar coefficients. In the following, we will focus on the question which structural properties the modified equation (2.5) shares with the original ODE (2.1).

Variational Structure

The ODE (2.1) has first order variational structure as it is the Euler–Lagrange equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{y}} - \frac{\partial L}{\partial y} = 0$$

to the variational principle

$$\delta S = 0 \quad \text{for} \quad S(y) = \int L(y(t), \dot{y}(t))\mathrm{d}t$$

with

$$L(y, \dot{y}) = \frac{1}{2}\|\dot{y}\|^2 + U(y).$$

Moreover, there exists a variational principle for (2.2):

Lemma 2.1. *Let the matrices $A_j \in \mathbb{R}^{n\times n}$ be symmetric. For $T > 0$ let $\mathbb{T}$ either be the circle $\mathbb{T} = \mathbb{R}/T\mathbb{Z}$ or the real line $\mathbb{T} = \mathbb{R}$. For y defined on $\mathbb{T}$ the variational principle*

$$\delta S_\Delta = 0 \quad \textit{for} \quad S_\Delta(y) = \int_{\mathbb{T}} L_\Delta\left(y(t), y(t+h), \dots, y\left(t + \frac{s}{2}h\right)\right)\mathrm{d}t \tag{2.6}$$

with

$$L_\Delta\left(y(t), y(t+h), \dots, y(t+\frac{s}{2}h)\right) = \frac{1}{2h^2}\left(\sum_{j=1}^{\frac{s}{2}} \langle A_j(y(t+jh) - y(t)), y(t+jh) - y(t)\rangle\right) + U(y(t)).$$

implies the functional equation (2.2)*. Moreover, if $\mathbb{T} = [a, b]$ is an interval, then* (2.6) *implies* (2.2) *on the interval $\mathring{\mathbb{T}} = [a + \frac{s}{2}h, b - \frac{s}{2}h]$. Here we assume that the function space for y and the potential U are such that $U \circ y$ and $\nabla U \circ y$ constitute square integrable functions.*

Proof. Let $\Delta_\tau y$ denote the forward difference, i.e. $(\Delta_\tau y)(t) = y(t+\tau) - y(t)$. Then $\langle y, \Delta_\tau z\rangle_{L^2(\mathbb{T},\mathbb{R}^n)} = \langle \Delta_\tau^* y, z\rangle_{L^2(\mathbb{T},\mathbb{R}^n)}$ holds with $\Delta_\tau^* = \Delta_{-\tau}$ on $\mathbb{T}$ if $\mathbb{T} \in \{\mathbb{R}, \mathbb{R}/T\mathbb{Z}\}$ or on the interval $[a+\tau, b-\tau]$ if $\mathbb{T} = [a, b]$. The expression $\Delta_\tau^*\Delta_\tau y$ corresponds to the central difference

$$(\Delta_\tau^*\Delta_\tau y)(t) = -y(t+\tau) + 2y(t) - y(t-\tau).$$

Let δ denote the variational derivative in the direction of a variation δy, i.e. $\delta S_\Delta(y) = \lim_{\epsilon\to 0}\frac{1}{\epsilon}(S_\Delta(y + \epsilon\delta y) - S_\Delta(y))$. Let $\delta y \in C_c^\infty(\mathbb{T}, \mathbb{R}^n)$, if $\mathbb{T} \in \{\mathbb{R}, \mathbb{R}/T\mathbb{Z}\}$ and let $\delta y \in C_c^\infty(\mathring{\mathbb{T}}, \mathbb{R}^n)$, if $\mathbb{T} = [a, b]$. We have

$$\delta S_\Delta(y) = \frac{1}{2h^2}\sum_{j=1}^{\frac{s}{2}} \delta\langle A_j\Delta_{jh}y, \Delta_{jh}y\rangle_{L^2(\mathbb{T},\mathbb{R}^n)} + \delta\int_{\mathbb{T}} U(y(t))\,\mathrm{d}t$$

$$= \frac{1}{h^2} \sum_{j=1}^{\frac{s}{2}} \langle A_j \Delta_{jh} y, \Delta_{jh} \delta y \rangle_{L^2(\mathbb{T},\mathbb{R}^n)} + \langle \nabla U(y), \delta y \rangle_{L^2(\mathbb{T},\mathbb{R}^n)}$$

$$= \frac{1}{h^2} \sum_{j=1}^{\frac{s}{2}} \langle A_j \Delta_{jh}^* \Delta_{jh} y + \nabla U(y), \delta y \rangle_{L^2(\mathbb{T},\mathbb{R}^n)}.$$

Now (2.2) follows from the fundamental lemma of the calculus of variations on $\mathbb{T}$ or $\mathring{\mathbb{T}}$, respectively.

To analyse structure preserving properties of the method (2.2), it might seem natural to seek a modified Lagrangian $L_{\mathrm{mod}}(y, \dot{y})$ given as a formal power series in the step-size h such that the modified variational principle

$$\delta S_{\mathrm{mod}} = 0 \quad \text{for} \quad S_{\mathrm{mod}}(y) = \int L_{\mathrm{mod}}(y(t), \dot{y}(t)) \mathrm{d}t$$

covers smooth solutions of (2.6) up to any order in the step-size h. However, we show that although a 1st order Lagrangian L_{mod} covering the modified equations always exists as a power series, it only exist in modified variables $(\tilde{y}, \dot{\tilde{y}})$ in the most general case. Even for simple methods, the existence of an expression in closed form for the change of coordinates from $(y, \dot{y})$ to $(\tilde{y}, \dot{\tilde{y}})$ can not be expected. This makes it difficult to compute L_{mod} using an ansatz.

Hamiltonian structure

Another approach is to work on the Hamiltonian side. The ODE (2.1) has the form of a Hamiltonian system

$$\dot{z}(t) = J^{-1} \nabla H(z(t)), \quad H(z) = \frac{1}{2} \|\dot{y}\|^2 - U(y), \quad z = \begin{pmatrix} y \\ \dot{y} \end{pmatrix}. \tag{2.7}$$

Here

$$J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} \tag{2.8}$$

is the standard symplectic structure.

In this paper we use a blended approach of the variational and Hamiltonian viewpoint to systematically compute a modified Hamiltonian system

$$\dot{z}(t) = J_{\mathrm{mod}}^{-1}(z(t)) \nabla H_{\mathrm{mod}}(z(t)) \tag{2.9}$$

consisting of a modified Hamiltonian H_{mod} and a modified symplectic structure J_{mod} given as formal power series in h such that for a truncation to arbitrary order N the ODE (2.9) covers the modified equation (2.5) up to higher order terms. Here J_{mod} is a skew symmetric matrix which satisfies a Jacobi identity. The following theorem can be considered as an instance of Theorem 1.1 and will be proved in Section 5.

Theorem 2.2. *Let $N \in \mathbb{N}$ denote the considered order of the series expansion of the matrix multistep method* (2.2)*, where the coefficients are symmetric matrices. There exists a modified symplectic structure $J_{\mathrm{mod}}^{[N]}$, which is $O(h)$ close to J and a modified Hamiltonian $H_{\mathrm{mod}}^{[N]}$, which is $O(h)$ close to H, such that*

$$\dot{z}(t) = (J_{\mathrm{mod}}^{[N]}(z))^{-1} \nabla H_{\mathrm{mod}}^{[N]}(z(t)) \tag{2.10}$$

with coordinate $z = (y, \dot{y})$ is equivalent to the modified equation (2.5) *up to terms of order $O(h^{N+1})$.*

Here $O(h)$-closeness of $J_{\mathrm{mod}}^{[N]}$ and J means that the zeroth coefficient of the polynomial $J_{\mathrm{mod}}^{[N]}$ in the formal variable h is given by J. $O(h)$-closeness of $H_{\mathrm{mod}}^{[N]}$ and H has an analogous meaning.

We observe conditions under which a modified first order Lagrangian $L_{\mathrm{mod}}(y, \dot{y})$ exists in the original variable y by analysing the modified symplectic structure J_{mod}.

Theorem 2.3. *If the matrices A_j in* (2.2) *are scalar multiples of the identity matrix, then there exists a modified Lagrangian $L_{\mathrm{mod}}^{[N]}$ depending on $(y, \dot{y})$ that is $O(h)$-close to $L(y, \dot{y})$ such that the modified variational principle*

$$\delta S_{\mathrm{mod}}^{[N]} = 0 \quad \textit{for} \quad S_{\mathrm{mod}}^{[N]}(y) = \int L_{\mathrm{mod}}^{[N]}(y(t), \dot{y}(t))\mathrm{d}t$$

yields the modified equation (2.5) *up to terms of order $O(h^{N+1})$.*

Again, $O(h)$-closeness is to be interpreted in a formal sense, analogously to its meaning in Theorem 2.2.

Theorem 2.3 applies to traditional multistep methods with scalar coefficients. However, we will see that for general linear multistep methods with matrix-valued coefficients the existence of a first order modified Lagrangian in the original variable y cannot be expected. A proof of Theorem 2.3 is postponed to Section 6.

3. Computation of modified hamiltonian structure

In this section we introduce a method to compute the modified data $J_{\mathrm{mod}}^{[N]}$ and $H_{\mathrm{mod}}^{[N]}$ of Theorem 2.2 such that (2.10) governs (2.5). We will then verify the validity of the construction method and prove Theorem 1.1 and 2.2 in Section 5.

Let $\mathcal{L}_\Delta^{[N]}$ denote the series expansion of

$$L_\Delta\left(y(t), y(t+h), \ldots, y\left(t + \frac{s}{2}h\right)\right)$$

to order N in the step-size h. The expression $\mathcal{L}_\Delta^{[N]}$ depends on the M-jet of y at t for some $M \in \mathbb{N}$. Notice that the order M variational principle

$$\delta \mathcal{S}_\Delta^{[N]}(y) = 0 \quad \text{with} \quad \mathcal{S}_\Delta^{[N]}(y) = \int \mathcal{L}_\Delta^{[N]}(y(t), \dot{y}(t), \ldots, y^{(M)}(t))\,\mathrm{d}t \tag{3.1}$$

recovers (2.4) up to higher order terms. We first compute a high-dimensional Hamiltonian system defined on the $2M-1$-jet space of y corresponding to the order M variational principle (3.1). The Hamiltonian principle is then reduced to a Hamiltonian system defined on the 1-jet space of y. It has the form (2.9) and covers (2.5) up to higher order terms.

To construct the high-dimensional Hamiltonian system, we use Ostrogradsky's Hamiltonian description of high-order Lagrangian systems [16]. For this we define variables

$$q = (y, \dot{y}, \ldots, y^{(M-1)})$$

and for $i = 1, \ldots, M$

$$p_i^j = \sum_{k=0}^{M-i} (-1)^k \frac{\mathrm{d}^k}{\mathrm{d}t^k} \frac{\partial \mathcal{L}_\Delta^{[N]}}{\partial (y_j)^{(k+i)}}.$$

Here the index j enumerates the components of y and $\frac{\mathrm{d}}{\mathrm{d}t}$ denotes the total derivative operator on the jet-space of y, which acts like

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho\left(y,\dot{y},\ldots,y^{(M)}\right)=\sum_{i=0}^{M}\left\langle\nabla_{y^{(i)}}\rho,y^{(i+1)}\right\rangle.$$

on a scalar valued function ρ defined on the M-jet space of y. The high dimensional Hamiltonian system consists of the Hamiltonian

$$\mathcal{H}^{[N]}=\sum_{i=1}^{M}\langle p_i,\dot{q}^i\rangle-\mathcal{L}_{\Delta}^{[N]},\tag{3.2}$$

where all expressions are expressed in $y^{[2M-1]}=(y,\dot{y},\ldots,y^{(2M-1)})$, and the symplectic structure matrix $\mathcal{J}_{\mathrm{mod}}^{[N]}(y^{[2M-1]})$. The skew-symmetric matrix $\mathcal{J}_{\mathrm{mod}}^{[N]}(y^{[2M-1]})$ is the representing matrix of the differential 2-form

$$\Omega^{[N]}=\sum_{i=1}^{M}\sum_{j=1}^{n}\mathrm{d}p_i^j\wedge\mathrm{d}q_j^i,\tag{3.3}$$

where p_i^j and q_j^i are interpreted as functions in the variable $y^{[2M-1]}$ of the $2M-1$-jet space, i.e. $\mathcal{J}_{\mathrm{mod}}^{[N]}$ is the anti-symmetrised tensor product* $\wedge$ of the gradients $\nabla_{y^{[2M-1]}}p_j^i$ and $\nabla_{y^{[2M-1]}}q_j^i$ summed over all indices.

To compute the modified Hamiltonian system on the 1-jet space with variable $y^{[1]}=(y,\dot{y})$, the variables $y^{(2)},\ldots,y^{(2M-1)}$ in the expression (3.2) for the Hamiltonian $\mathcal{H}^{[N]}(y^{[2M-1]})$ are repeatedly replaced by (2.5) until higher derivatives only occur in $O(h^{N+1})$ terms. This yields $H_{\mathrm{mod}}^{[N]}(y,\dot{y})$. Similarly, we can consider p_i^j and q_j^i as functions of $(y,\dot{y})$ truncating $O(h^{N+1})$ terms. The matrix $J_{\mathrm{mod}}^{[N]}$ is then given as the representing matrix of the 2-form $\Omega^{[N]}$ pulled to the 1-jet space of the variable y, i.e. interpreting $y^{[1]}=(y,\dot{y})$ as the only independent variables. Equivalently, $J_{\mathrm{mod}}^{[N]}$ is the anti-symmetrised tensor product $\wedge$ of the gradients $\nabla_{y^{[1]}}p_j^i$ and $\nabla_{y^{[1]}}q_j^i$ summed over all indices. (As $J_{\mathrm{mod}}^{[N]}$ is constructed from a closed differential 2-form, it is automatically skew-symmetric and satisfies the Jacobi identity.) This completes the construction of the modified data.

The system (2.9) recovers (2.5) up to higher order terms as we will prove after a computational example.

4. Computational example

As introduced in Section 2, consider the multistep method

$$A_2y(t-2h)+A_1y(t-h)-2(A_1+A_2)y(t)+A_1y(t+h)+A_2y(t+2h)=h^2\nabla U(y(t))\tag{4.1}$$

with matrix coefficients in dimension $n=2$. By the consistency requirement, $A_2=1/4(I-A_1)$. We obtain

$$J_{\mathrm{mod}}{}^{[4]}=J+h^2\begin{pmatrix}J_{11}^2&-J_{21}^2\\J_{21}^2&0\end{pmatrix}+h^4\begin{pmatrix}J_{11}^4&-J_{21}^4\\J_{21}^4&J_{22}^4\end{pmatrix}$$

with

$$J_{11}^2=\begin{pmatrix}0&-b_1\\b_1&0\end{pmatrix}$$

*This corresponds to the command `TensorWedge` in Wolfram Mathematica.

where

$$b_1 = \frac{1}{4}\Big(\dot{y}_2\left(a_{12}(U^{(2,1)} - U^{(0,3)}) + (a_{22} - a_{11})\,U^{(1,2)}\right) + \quad \dot{y}_1\left(-a_{12}(U^{(2,1)} - U^{(0,3)}) + (a_{22} - a_{11})\,U^{(2,1)}\right)\Big).$$

Here and in the following $U^{(k,l)} = \frac{\partial^{k+l}U(y)}{\partial^k y_1 \partial^l y_2}$. As the expressions of higher order terms become quite complicated, we refer the reader to the Mathematica Notebooks of our accompanying source code [11]. However, as this will be relevant in the discussion later, we are reporting J^4_{22} for the special case that A_1 is a diagonal matrix: we have

$$J^4_{22} = \begin{pmatrix} 0 & -b_2 \\ b_2 & 0 \end{pmatrix}$$

with

$$b_2 = \frac{1}{8}(a_{11}^2 - a_{22}^2 + 2(a_{22} - a_{11}))(U^{(2,1)}\dot{y}_1 + U^{(1,2)}\dot{y}_2) \quad \text{if} \quad A_1 = \begin{pmatrix} a_{11} & 0 \\ 0 & a_{22} \end{pmatrix}.$$

The modified Hamiltonian $H^{[4]}_{\text{mod}}$ for the general case is given as

$$H^{[4]}_{\text{mod}}(y, \dot{y}) = H(y, \dot{y}) + h^2 H_2(y, \dot{y}) + h^4 H_4(y, \dot{y})$$

with

$$\begin{aligned} H_2(y, \dot{y}) = \frac{1}{24}\Big(&(\dot{y}_1)^2(2(4 - 3a_{11})U^{(2,0)} - 6a_{12}U^{(1,1)}) \\ &- 2\dot{y}_2\dot{y}_1(3a_{12}U^{(0,2)} + (3a_{11} + 3a_{22} - 8)U^{(1,1)} + 3a_{12}U^{(2,0)}) \\ &- 6a_{22}U^{(0,2)}(\dot{y}_2)^2 - 6a_{12}U^{(1,1)}(\dot{y}_2)^2 \\ &+ (3a_{22} - 4)(U^{(0,1)})^2 + 3a_{11}(U^{(1,0)})^2 \\ &+ 6a_{12}U^{(0,1)}U^{(1,0)} + 8U^{(0,2)}(\dot{y}_2)^2 - 4(U^{(1,0)})^2\Big). \end{aligned}$$

For further terms, we refer the reader to the Mathematica Notebooks of our accompanying source code. Hamilton's equations

$$\dot{z} = (J^{[4]}_{\text{mod}}(z))^{-1}\nabla H^{[4]}_{\text{mod}}(z)$$

for the modified Hamiltonian system are equivalent to the modified equation (2.5) truncating terms of order $O(h^6)$.

Figure 1 shows a numerical experiment with a rotational invariant potential. The start values for the multistep formula were obtained using the fourth order modified equation. Trajectories computed with the multistep scheme look very regular. The quantities $H^{[0]}_{\text{mod}} = H$, $H^{[2]}_{\text{mod}}$, and $H^{[4]}_{\text{mod}}$ evaluated along a trajectory show oscillatory energy error behaviour. Experiments with different values for the step-size h confirm the preservation of $H^{[k]}_{\text{mod}}$, $k = 0, 2, 4$ up to truncation error. Initialising the multistep scheme with the fourth order modified equation, the effects of spurious solutions was minimised. However, as h is decreased, spurious solutions cause wriggles in the energy error $H^{[k]}_{\text{mod}} - H(z_{\text{init}})$, $k = 0, 2, 4$ which eventually prevent further energy error decay.

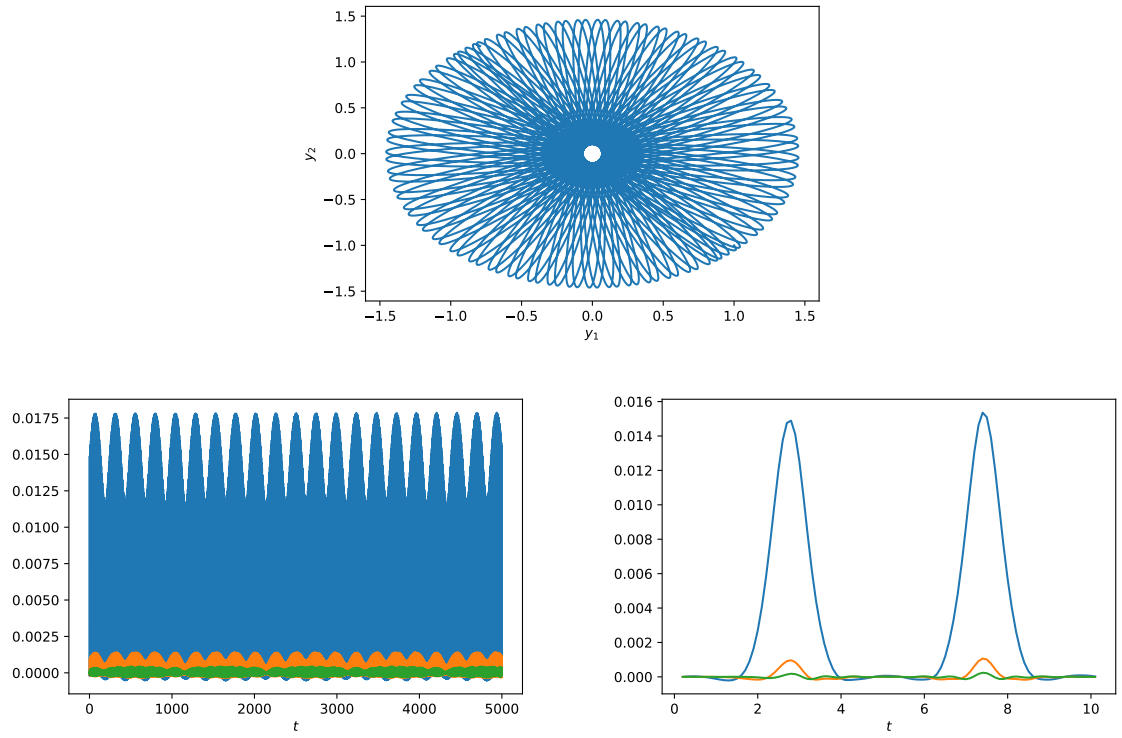

Figure 1. Numerical experiment with the multistep scheme (4.1) with $A_2 = \mathrm{diag}\,(0.85, 1.25)$, $A_1 = I - 4A_2$, $U(y) = \exp\left(-\frac{1}{2}(y_1^2 + y_2^2)\right)$ and time-step $h = 0.1$. The multistep formula was initialised at times 0, h, $2h$, $3h$ by integrating the order 4 modified equation using very fine Euler-steps starting from the initial value $(y_{\mathrm{init}}, \dot{y}_{\mathrm{init}}) = ((1, -1), (0.1, -0.2))$. The figure at the top shows a trajectory up to time $t = 500$, which looks like an orbit in a completely integrable system. The figures below show an evaluation along the trajectory of $H - H((y_{\mathrm{init}}, \dot{y}_{\mathrm{init}}))$ (blue) as well as $H^{[2]}_{\mathrm{mod}} - H^{[2]}_{\mathrm{mod}}((y_{\mathrm{init}}, \dot{y}_{\mathrm{init}}))$ (orange) and $H^{[4]}_{\mathrm{mod}} - H^{[4]}_{\mathrm{mod}}((y_{\mathrm{init}}, \dot{y}_{\mathrm{init}}))$ (green) up to time $t = 5000$ and $t = 10$, respectively. We see oscillatory energy error behaviour.

If $A_2 = \alpha I$ with $\alpha \geq \frac{1}{4}$ and $A_1 = I - 4A_2$, then (4.1) corresponds to a classical stable[†] multistep scheme: the generating polynomial ρ to (4.1) is given as

$$\rho(\xi) = \alpha + (1 - 4\alpha)\xi - 2(1 - 3\alpha)\xi^2 + (1 - 4\alpha)\xi^3 + \alpha\xi^4.$$

The polynomial ρ has a double root at 1 as well as the roots

$$\xi_3 = 1 - \frac{1}{2\alpha} + \frac{\sqrt{1 - 4\alpha}}{2\alpha}, \quad \xi_4 = 1 - \frac{1}{2\alpha} - \frac{\sqrt{1 - 4\alpha}}{2\alpha}.$$

Since $\alpha \geq \frac{1}{4}$, the roots ξ_3 and ξ_4 are complex conjugate to each other and lie on the unit circle such that the scheme is stable.

[†]A multistep scheme for second order equations is stable if all roots of its generating polynomial lie in the closed unit disk and those on the circle are at most double zeros [3, XV.1.2].

Moreover, L_Δ and $\mathcal{L}_\Delta$ are rotationally invariant because A_1 and A_2 commute with rotation matrices. An application of Noether's theorem to $\mathcal{L}_\Delta^{[N]}$ yields the following modified angular momentum:

$$\mathcal{I}^{[N]} = \sum_{m=1}^{M} \sum_{k=0}^{m-1} (-1)^k \left\langle \nabla_{y^k} \mathcal{L}_\Delta^{[N]}, \mathrm{d}R y^{(m-1-k)} \right\rangle, \quad \mathrm{d}R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

The integer M is the order of the highest derivative of y in $\mathcal{L}_\Delta^{[N]}$. For the truncation order $N = 4$ we have $M = 5$. Using (2.5) repeatedly, derivatives of y of order greater than two are replaced by terms in y, $\dot{y}$ and $O(h^6)$-terms. After truncation at order 4 in h we obtain the modified angular momentum $I_{\mathrm{mod}}^{[4]}(y, \dot{y})$. Figure 2 shows the conservation error of $I = I_{\mathrm{mod}}^{[0]}, I_{\mathrm{mod}}^{[2]}, I_{\mathrm{mod}}^{[4]}$ along a trajectory. We see oscillatory error behaviour. The oscillations shrink as the step-size h decreases until spurious solutions prevent further decrease.

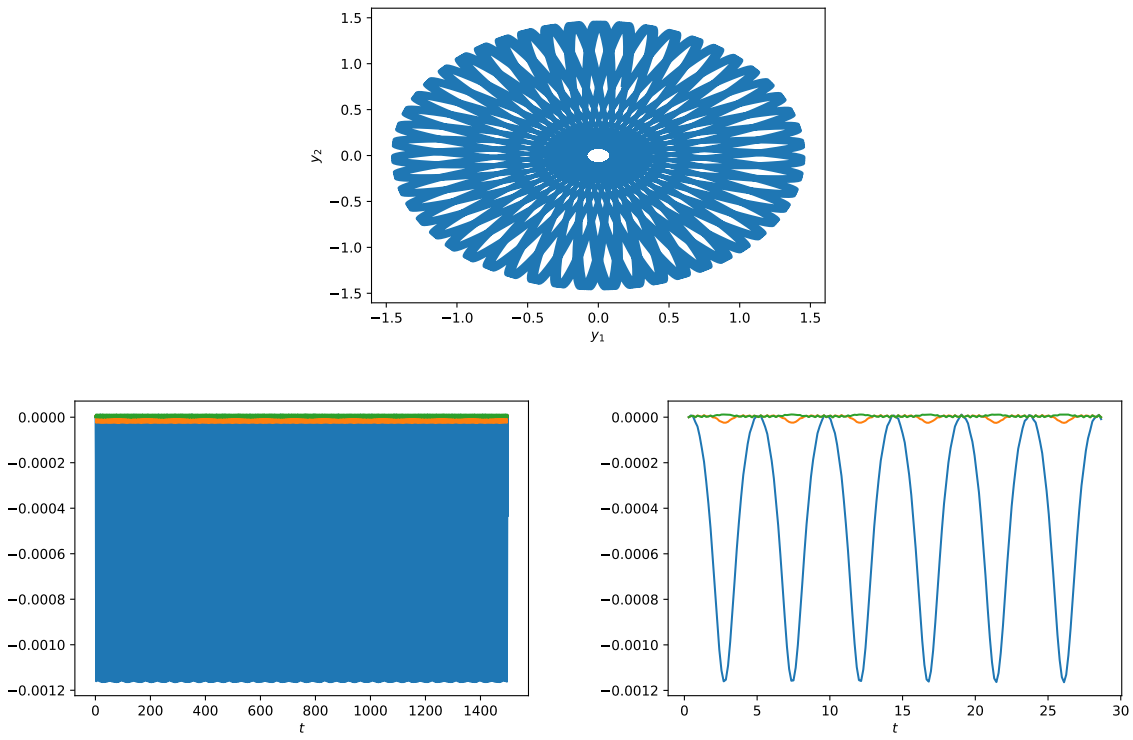

Figure 2. The plots in the second row show the error $I - I(y_{\mathrm{init}}, \dot{y}_{\mathrm{init}})$ (blue) in the angular momentum and $I_{\mathrm{mod}}^{[2]} - I_{\mathrm{mod}}^{[2]}(y_{\mathrm{init}}, \dot{y}_{\mathrm{init}})$ (orange) and $I_{\mathrm{mod}}^{[4]} - I_{\mathrm{mod}}^{[4]}(y_{\mathrm{init}}, \dot{y}_{\mathrm{init}})$ (green) along a trajectory. The same data as in Figure 1 was used apart from $A_2 = \mathrm{diag}(0.3, 0.3)$ and $h = 0.15$. The plot at the top shows a trajectory initialised as in Figure 1.

For source code of our experiments refer to [11].

5. Validity of the construction method

Let us prepare the proof of Theorem 1.1. We will use the language of Differential Geometry.

To motivate the following proposition and to fix notation recall the following fact: if (Z', ω', H') is a Hamiltonian system with Hamiltonian vector field $X'_{H'}$, $Z \overset{\Psi}{\hookrightarrow} Z'$ an embedding such that $\Psi(Z)$ is a symplectic submanifold that is invariant under motions, i.e. $X'_{H'}(\Psi(z)) \in T_{\Psi(z)}\Psi(Z)$ for all $z \in Z$, then the pull-back system (Z, ω, H) with $\omega = \Psi^*\omega'$, $H = H' \circ \Psi$ is a Hamiltonian system with Hamiltonian vector field X_H. Here $\Psi^*\omega'$ denotes the pull-back of ω' along Ψ. The Hamiltonian vector fields X_H and $X'_{H'}$ relate by

$$\Psi_* X_H = X'_{H'} \circ \Psi,$$

i.e. for a motion γ with $\dot\gamma = X_H \circ \gamma$ the curve $\gamma' = \Psi \circ \gamma$ is a motion of (Z', ω', H'), i.e. $\dot\gamma' = X'_{H'} \circ \gamma'$. Here Ψ_* denotes the push-forward map, i.e. $(\Psi_* X_H)(z) = \mathrm{d}\Psi|_z(X_H(z))$ for $z \in Z$.

In the following, we will adapt this statement to a setting, where the definition of ω', H', and Ψ' contain a formal variable h and where $\Psi(Z)$ is left invariant only up to higher order terms in h.

Definition 5.1 (Formal Hamiltonian system)**.** Let $N \in \mathbb{N}$ be a truncation index, Z be a smooth manifold, h a formal variable, and $H = \sum_{k=0}^{N} h^k H_k$ a formal polynomial whose coefficients are smooth maps $Z \to \mathbb{R}$. Further, let $\omega = \sum_{k=0}^{N} h^k \omega_k$ be a formal polynomial whose coefficients are 2-forms on Z. The collection (Z, ω, H) is called a *formal Hamiltonian system with truncation index N* if the following conditions are satisfied.

- The formal symplectic form ω is closed, i.e. $\mathrm{d}\omega = \sum_{k=0}^{N} h^k \mathrm{d}\omega_k$ is a formal polynomial whose coefficients are 3-forms that are zero.
- The 2-form ω_0 is non-degenerate.

The relation $-\mathrm{d}H = \omega(X, \cdot)$ defines a formal power series $X = \sum_{k=0}^{\infty} h^k X_k$, where X_k are vector fields on Z. The truncation $X_H := \sum_{k=0}^{N} h^k X_k$ is called *(formal) Hamiltonian vector field.* The formal differential equation $\dot\gamma = X_H \circ \gamma$ is called *Hamilton's equation*.

Proposition 5.2. *Let $N \in \mathbb{N}$ be a truncation index, Z' and Z be manifolds, and let (Z', ω', H') be a formal Hamiltonian system with truncation index N. Consider a formal polynomial $\Psi = \sum_{k=0}^{N} \Psi_k h^k$ such that $\Psi_k \colon Z \to Z'$ are smooth and $\Psi_0 \colon Z \to Z'$ is an embedding. To $z \in Z$ define formal tangent spaces*

$$T_{\Psi(z)}\Psi(Z) := \left\{ \sum_{k=0}^{N} h^k \mathrm{d}\Psi_k|_z(v) \,\middle|\, v \in T_z Z \right\} \subset \bigoplus_{k=0}^{N} h^k T_{z'_k} Z',$$

where $z' := \sum_{k=0}^{N} h^k z'_k := \Psi(z)$.

- *Assume that the Hamiltonian vector field is tangential to $\Psi(Z)$, i.e. $\mathrm{trk}(X'_{H'}(\Psi(z))) \in T_{\Psi(z)}\Psi(Z)$ for all $z \in Z$. Here* trk *denotes the truncation of $O(h^{N+1})$-terms.*
- *Assume that $\Psi(Z)$ is a symplectic submanifold of Z', i.e. for all $z \in Z$*

$$T_{\Psi(z)}\Psi(Z) \cap T_{\Psi(z)}\Psi(Z)^{\perp_{\omega'}} = 0.$$

Here $\Psi(Z)^{\perp_{\omega'}}$ denotes the symplectic complement

$$T_{z'}\Psi(Z)^{\perp_{\omega'}} = \left\{ v \in \bigoplus_{k=0}^{N} h^k T_{z'_k} Z' \,\middle|\, \mathrm{trk}(\omega'(v, w)) = 0 \ \forall w \in T_{z'}\Psi(Z) \right\},$$

where $z' := \sum_{k=0}^{N} h^k z'_k := \Psi(z)$.

Then the pull-back system (Z, ω, H) with $\omega = \Psi^*\omega'$, $H = H' \circ \Psi$ constitutes a formal Hamiltonian system with truncation index N and

$$\operatorname{trk}(\Psi_* X_H) = X'_{H'} \circ \Psi.$$

Proof. Let $\omega = \sum_{k=0}^{N} h^k \omega_k$ Since Ψ_0 is an embedding, it is an immersion and we conclude that ω_0 is non-degenerate as ω'_0 is non-degenerate. Since pull-back and the differential d commute, ω is closed.

In the following calculations we identify two formal polynomials P_1 and P_2 with coefficients of the same type (real numbers, n-forms, smooth functions, ...) if and only if $P_1 - P_2 \in O(h^{N+1})$. For all $z \in Z$ we have

$$\begin{aligned} -\omega'_{\Psi(z)}(\mathrm{d}\Psi|_z(X_H(z)), \mathrm{d}\Psi|_z(\cdot)) &= -(\Psi^*\omega')_z(X_H(z), \cdot) = -\omega_z(X_H(z), \cdot) \\ &= \mathrm{d}H|_z = \mathrm{d}(H' \circ \Psi)|_z = \mathrm{d}H'|_{\Psi(z)} \circ \mathrm{d}\Psi|_z \\ &= -\omega'_{\Psi(z)}(X'_{H'}(\Psi(z)), \mathrm{d}\Psi|_z(\cdot)) \end{aligned}$$

$$\begin{aligned} &\implies \omega'_{\Psi(z)}(\mathrm{d}\Psi|_z(X_H(z)) - X'_{H'}(\Psi(z)), \mathrm{d}\Psi|_z(\cdot)) = 0 \\ &\implies \mathrm{d}\Psi|_z(X_H(z)) - X'_{H'}(\Psi(z)) \in T_{\Psi(z)}\Psi(Z)^{\perp_{\omega'}}. \end{aligned}$$

As the inclusion $\mathrm{d}\Psi|_z(X_H(z)) - X'_{H'}(\Psi(z)) \in T_{\Psi(z)}\Psi(Z)$ holds as well and $\Psi(Z)$ is a symplectic submanifold, it follows that

$$\mathrm{d}\Psi|_z(X_H(z)) = X'_{H'}(\Psi(z)).$$

We can now proceed to the proof of Theorem 1.1.

Proof. Step 1. Construction and validity of the modified equation. The truncated power series $\mathcal{L}_\Delta^{[N]}$ is a formal polynomial of the form

$$\mathcal{L}_\Delta^{[N]}(y^{[M]}) = L(y, \dot{y}) + \sum_{k=1}^{N} h^k \mathcal{L}_\Delta^k(y^{[M]}),$$

where h is the formal variable. Since the Lagrangian L is regular, the Euler–Lagrange equations to $\mathcal{L}_\Delta^{[N]}$

$$\sum_{j=0}^{M} (-1)^j \frac{\mathrm{d}}{\mathrm{d}t}\left(\nabla_{y^{(j)}} \mathcal{L}_\Delta^{[N]}(y^{[M]})\right) = 0$$

are equivalent to an ordinary differential equation of the form

$$\ddot{y} = g_0(y, \dot{y}) + \sum_{k=1}^{N} h^k \tilde{g}_k(y^{[2M]}) \tag{5.1}$$

when truncating terms of order $O(h^{[N+1]})$. The ordinary differential equation is of order $2M$ because $\mathcal{L}_\Delta^{[N]}(y^{[M]})$ is regular as well. Iterative replacements of derivatives of order $j \geq 2$ by derivatives of (5.1) yield the modified equation

$$\ddot{y} = g_0(y, \dot{y}) + \sum_{k=1}^{N} h^k g_k(y, \dot{y}) + O(h^{[N+1]}). \tag{5.2}$$

Solutions to

$$\ddot{y} = g_0(y, \dot{y}) + \sum_{k=1}^{N} h^k g_k(y^{[M]}) \tag{5.3}$$

fulfil (5.1) up to $O(h^{[N+1]})$ terms by construction. This proves the first part of Theorem 1.1.

Step 2. Existence of Hamiltonian structure on a higher jet-space. Let Y denote the domain of y (smooth manifold), $Z = \mathrm{Jet}^1(Y)$ the 1-jet space, and $Z' = \mathrm{Jet}^{2M-1}(Y)$. The iterative substitution procedure gives rise to a formal polynomial Ψ of maps $Z \to Z'$ defined by $y^{[1]} \mapsto y^{[2M-1]}$, where $y^{(j)}$ for $j \geq 2$ is expressed as a formal polynomial of functions depending on $y^{[1]} = (y, \dot{y})$. The expression for $y^{(j)}$ is obtained by deriving (5.1) $j-2$ times, iteratively replacing derivatives of order greater than 2 by derivatives of (5.1) followed by a truncation of $O(h^{[N+1]})$ terms.

In the remainder of this part of the proof we suppress the fact that we operate on formal polynomials as the following steps can also be done when h is substituted with a sufficiently small real number $h > 0$.

In the following, we will show that there exists a Hamiltonian structure (Z', ω', H') on Z' whose motions correspond to the motions induced by the order M-Lagrangian $\mathcal{L}_\Delta^{[N]}(y^{[M]})$. To compute the Hamiltonian structure (Z', ω', H'), we first construct a Hamiltonian system on $T^*\mathrm{Jet}^{M-1}(Y)$ which we will then pull back to Z'.

As the Lagrangian $\mathcal{L}_\Delta^{[N]}(y^{[M]})$ is regular, by Ostrogradsky's principle for high order Lagrangians [16], there exists a transformation $\chi\colon Z' \to T^*\mathrm{Jet}^{M-1}(Y)$ between the jet variable $y^{[2M-1]} \in Z'$ and variables $(q, p) \in T^*\mathrm{Jet}^{M-1}(Y)$ with

$$q = (y, \dot{y}, \ldots, y^{(M-1)})$$

and

$$p_i = \sum_{k=0}^{M-i} (-1)^k \frac{\mathrm{d}^k}{\mathrm{d}t^k} \nabla_{y^{(k+i)}} \mathcal{L}_\Delta^{[N]}(y^{[M]}), \quad i = 1, \ldots, M$$

such that with

$$\Omega = \sum_{i=1}^{M} \sum_{j=1}^{n} \mathrm{d}p_i^j \wedge \mathrm{d}q_j^i, \quad \text{and} \quad \mathcal{H}^{[N]}(q, p) = \sum_{k=1}^{M} \langle p_k, q^k \rangle - \mathcal{L}_\Delta^{[N]}$$

the motions of (5.1) are exactly mapped to the motions of the Hamiltonian vector field $X_{\mathcal{H}^{[N]}}$ on $T^*\mathrm{Jet}^{M-1}(Y)$ with

$$-\mathrm{d}\mathcal{H}^{[N]} = \Omega(X_{\mathcal{H}^{[N]}}, \cdot) \tag{5.4}$$

by the transformation χ^{-1}. Let $\omega' = \chi^*\Omega$ and $H' = \mathcal{H}^{[N]} \circ \chi$. Now (Z', ω', H') is a Hamiltonian system on Z' whose motions are exactly the motions induced by the Lagrangian structure. This completes step 2 of the proof.

Remark 5.3. In the setting of formal polynomials, Ω is a two form whose coefficients are formal polynomials or, alternatively, Ω is a formal polynomial whose coefficients are 2-forms. In this case, (5.4) defines a formal series $X_{\mathcal{H}^{[N]}}$ in h from which we truncate $O(h^{N+1})$ terms. This is also done when defining $X'_{H'}$ through $-\mathrm{d}H' = \omega'(X'_{H'}, \cdot)$. The differential equation defined by $X'_{H'}$ is then equivalent to the differential equation induced by the Lagrangian structure up to $O(h^{N+1})$ terms.

Step 3. Pull-back of (Z', ω', H') *to* Z. As the Lagrangian L is regular, the 2-form $\omega'_0 = \sum_{j=1}^{n} \mathrm{d}y_j \wedge \mathrm{d}\mathfrak{p}^j$, where $\mathfrak{p}$ is obtained by the Legendre transform for L, is a symplectic form. The pull-back form $\omega = \Psi^*\omega'$ is a formal polynomial in h, whose coefficients are 2-forms. The zeroth coefficient ω_0 coincides with ω'_0. ω is, therefore, non-degenerate and $\Psi(Z)$ is a symplectic submanifold of Z' in the sense of Proposition 5.2. Moreover, by construction of H' and Ψ the condition $\mathrm{trk}(X'_{H'}(\Psi(z))) \in T_{\Psi(z)}\Psi(Z)$ for all $z \in Z$ is fulfilled. Therefore, Proposition 5.2 completes the theorem's proof.

Proof of Theorem 2.2. The construction method of the modified data $J^{[N]}_{\mathrm{mod}}$ and $H^{[N]}_{\mathrm{mod}}$ coincides with the construction method verified in the proof of Theorem 1.1.

6. On the existence of modified lagrangians

Proposition 6.1. *All Hamiltonian systems*

$$\dot{z}(t) = \bar{J}^{-1}(z(t))\nabla\bar{H}(z(t)), \quad z = \begin{pmatrix} y \\ \dot{y} \end{pmatrix}$$

with a fixed symplectic structure represented by $\bar{J}$ *can be formulated as variational problems of the form*

$$\delta\bar{S} = 0 \quad \textit{for} \quad \bar{S}(y) = \int \bar{L}(y(t), \dot{y}(t))\mathrm{d}t \tag{6.1}$$

if and only if $\bar{J}$ *is of the form*

$$\bar{J} = \begin{pmatrix} * & * \\ * & 0_n \end{pmatrix}, \tag{6.2}$$

where 0_n *denotes an* $n \times n$*-dimensional zero matrix with* n *the dimension of* y*, (i.e. if the distribution spanned by the vector fields* $\frac{\partial}{\partial\dot{y}_1}, \ldots, \frac{\partial}{\partial\dot{y}_n}$ *is Lagrangian).*

Proof. Denote the domain of the variable y by Y, the 1-jet space over Y by $\mathrm{Jet}^1(Y)$, and the symplectic form represented by the matrix $\bar{J}$ by $\bar{\omega}$. Consider a Hamiltonian $\bar{H}\colon \mathrm{Jet}^1(Y) \to \mathbb{R}$. If the distribution $\mathcal{D}$ spanned by $\frac{\partial}{\partial\dot{y}_1}, \ldots, \frac{\partial}{\partial\dot{y}_n}$ is Lagrangian w.r.t. $\bar{\omega}$, then there exists a primitive $\bar{\lambda}$ of $\bar{\omega}$ with kernel $\mathcal{D}$ [7, Cor.15.7]. The 1-form $\bar{\lambda}$ is of the form

$$\bar{\lambda} = \sum_{j=1}^{n} \ell(y, \dot{y})\,\mathrm{d}y_j.$$

By Hamilton's principle, a curve $\gamma\colon [t_{\mathrm{start}}, t_{\mathrm{end}}] \to Y$ is a Hamiltonian motion if the action functional

$$\bar{S}(\gamma) = \int_\gamma (\bar{\lambda} - \bar{H}\mathrm{d}t)$$

is stationary w.r.t. variations of γ through smooth curves fixing the endpoints. Thanks to the special structure of $\bar{\lambda}$ (absence of $\mathrm{d}\dot{y}_j$-terms), the pullback of $\bar{\lambda} - \bar{H}\mathrm{d}t$ along a curve γ has the form

$$\bar{L}(y(t), \dot{y}(t))\mathrm{d}t$$

with $y(t)$ describing $\gamma(t)$ in the coordinate y. Therefore, in coordinates, the variational principle has the form (6.1).

On the other hand, a variational principle of the form (6.1) with regular Lagrangian $\bar{L}$ (i.e. invertible $\left(\frac{\partial^2 \bar{L}}{\partial \dot{y}_i \partial \dot{y}_j}\right)_{i,j}$) can be formulated as a Hamiltonian system with Hamiltonian

$$\bar{H}(y, \dot{y}) = \langle \dot{y}, p(y, \dot{y}) \rangle - L(y, \dot{y}), \quad \text{and with} \quad p(y, \dot{y}) = \frac{\partial \bar{L}}{\partial \dot{y}}(y, \dot{y}).$$

The symplectic structure is given as

$$\bar{\omega} = \mathrm{d}\bar{\lambda}, \quad \text{with} \quad \bar{\lambda} = \sum_{j=1}^{n} p_j(y, \dot{y}) \mathrm{d}y_j.$$

As the distribution $\mathcal{D}$ spanned by $\frac{\partial}{\partial \dot{y}_1}, \ldots, \frac{\partial}{\partial \dot{y}_n}$ is in the kernel of a primitive of $\bar{\omega}$, the distribution is Lagrangian.

Remark 6.2. The strength of Proposition 6.1 lies in the assertion that $\bar{L}$ is a first-order Lagrangian in the original variable y, i.e. it depends on $(y, \dot{y})$ only. If $\bar{J}$ is *not* of the required form, then, by Darboux's theorem, we can perform a change of variables on $\mathrm{Jet}^1(Y)$ such that $\bar{\omega}$ is the standard symplectic form $\sum \mathrm{d}\bar{p}_i \wedge \mathrm{d}\bar{q}_i$ and $\bar{L}$ is a 1st-order Lagrangian in q, i.e. depends on $(q, \dot{q})$ but not on higher derivatives in q. However, as q and p each depend on $(y, \dot{y})$, the variables have lost their dynamical meaning. This is because the required change of variables on $\mathrm{Jet}^1(Y)$ is not fibred, i.e. the jet-space structure is not preserved. Expressed in the original variable y, the Lagrangian $\bar{L}$ then depends on $(y, \dot{y}, \ddot{y})$, i.e. describes a higher order variational structure.

Remark 6.3. The computational example presented in Section 4 provides an example for which the modified symplectic structure is *not* of the form that is required in Proposition 6.1, unless A_1 is of the form $A_1 = \alpha I$, i.e. the method coincides with a multistep method with scalar coefficients. Indeed, if the considered method is a classical multistep method, i.e. all coefficients are scalar, then $L_{\mathrm{mod}}(y, \dot{y})$ exists by Theorem 2.3.

We now proceed to the proof of Theorem 2.3. We exploit that linear multistep methods can be interpreted as 1-step methods on the original phase space [3, §XV.2]. Here and below we refer to the theory of linear multistep methods for 2nd order ODEs.

Proof of Theorem 2.3. As proved in [1, §5], for the underlying 1-step method ϕ of a symmetric linear multistep method there exists a local diffeomorphism ψ such that $\tilde{\phi} = \psi \circ \phi \circ \psi^{-1}$ is symplectic with respect to the original symplectic structure $\omega = \sum_j \mathrm{d}p_j \wedge \mathrm{d}q_j$. The conjugacy ψ is given as a P-series applied to the original Hamiltonian vector field X_0 given by the right hand side of

$$\begin{aligned} \dot{q} &= p \\ \dot{p} &= \nabla U(q). \end{aligned}$$

The P-series ψ is a formal power series that is in general not convergent. Conjugacy, pull-back and push-forward operations are to be interpreted in a formal sense. The map ϕ is symplectic with respect to the modified symplectic structure $\omega_{\mathrm{mod}} = \psi^* \omega$ and is the time-h-flow of a vector field X, for which the flow

equations correspond to a first order formulation of the modified equation (2.5). The ω-symplectic map $\tilde{\phi}$ is the time-h-flow of the ψ-related vector field $\tilde{X} = \psi_* X \circ \psi^{-1}$. By standard results on backward error analysis for symplectic integrators [3, §IX], $\tilde{X}$ is a Hamiltonian vector field w.r.t. the standard symplectic structure ω for a Hamiltonian H up to any order in the step-size h. Pulling back the Hamiltonian structure (ω, H) using ψ, we obtain a modified Hamiltonian system $(\omega_{\text{mod}}, H_{\text{mod}}) = (\psi^*\omega, H \circ \psi)$ such that Hamilton's equations are equivalent to the modified equation (2.5).

Since ψ is a P-series in X_0, the distribution $\mathcal{D}$ spanned by the vertical vector fields $\frac{\partial}{\partial p_1}, \ldots, \frac{\partial}{\partial p_n}$ is Lagrangian w.r.t. $\omega_{\text{mod}} = \psi^*\omega$. By Proposition 6.1, for any order N in the step-size h there exists a first order Lagrangian $L_{\text{mod}}(y, \dot{y})$ in the original variable such that the variational principle

$$\delta\left(\int L_{\text{mod}}(y(t), \dot{y}(t))\mathrm{d}t\right) = 0$$

recovers the modified equation up to higher order terms in h.

Remark 6.4. The proof of Theorem 2.3 also shows that L_{mod} in Theorem 2.3 has the structure of an S-series applied to a P-series (see [1]). The modified Lagrangian L_{mod} can, thus, be computed with an ansatz as well. The modified data H_{mod} and J_{mod} can then be computed from L_{mod} by a Legendre transformation.

7. Future work

Motivated by optimal truncation results for modified equations, it would be interesting to analyse the convergence properties of modified symplectic structures, modified Hamiltonians, and modified Lagrangians. Moreover, in view of Remark 6.4, a systematic description of the combinatorial structure of the modified quantities appears feasible.

Acknowledgements

We would like to thank Peter Hydon and Chris Budd for discussions on multisymplecticity and variational principles at the Isaac Newton Institute of the University of Cambridge in 2019. This research was supported by the Marsden Fund of the Royal Society Te Apārangi and the School of Fundamental Sciences, Massey University, Manawatū, New Zealand. The authors would like to thank the Isaac Newton Institute for Mathematical Sciences for support and hospitality during the programme *Geometry, compatibility and structure preservation in computational differential equations* (2019) when work on this paper was undertaken. RM would like to thank the Simons Foundation for their support. This work was supported by: EPSRC grant number EP/R014604/1. The authors acknowledge support from the European Union Horizon 2020 research and innovation programmes under the Marie Skodowska-Curie grant agreement No. 691070.

References

1. P. Chartier, E. Faou, A. Murua. An algebraic approach to invariant preserving integators: The case of quadratic and Hamiltonian invariants. *Numer. Math*, **103** (2006), 575–590. https://doi.org/10.1007/s00211-006-0003-8

2. C. L. Ellison, J. W. Burby, J. M. Finn, H. Qin, W. M. Tang, Initializing and stabilizing variational multistep algorithms for modeling dynamical systems, 2014, *arXiv preprint arXiv*.
3. E. Hairer, C. Lubich, G. Wanner. Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations. Springer Series in Com- putational Mathematics, 2013 https://doi.org/10.1007/3-540-30666-8
4. L. Jódar, J. L. Morera, E. Navarro, On convergent linear multistep matrix methods, *Int. J. Comput. Math*, **40** (1991), 211–219. https://doi.org/10.1080/00207169108804014
5. J. D. Lambert, S. Sigurdsson, Multistep methods with variable matrix coefficients, *Siam. J. Numer. Anal*, **9** (1972), 715–733. https://doi.org/10.1137/0709060
6. B. Leimkuhler, S. Reich, Simulating Hamiltonian Dynamics, *Cambridge University Press*, **2** (2005). https://doi.org/10.1017/CBO9780511614118
7. P. Libermann, C.M. Marle, Symplectic manifolds and Poisson manifolds, *In Symplectic Geometry and Analytical Mechanics.*, (1987) 89–184. https://doi.org/10.1007/978-94-009-3807-6-3
8. J.E. Marsden, M. West. Discrete mechanics and variational integrators. *Acta Numerica*, **10** (2001), 357–514. https://doi.org/10.1017/S096249290100006X
9. R. I. McLachlan, C. Offen, Backward error analysis for variational discretisa-tions of PDEs, *J. Geom. Mech*, **14** (2022), 447. https://doi.org/10.3934/jgm.2022014
10. S. Ober-Blöbaum, C. Offen, Variational learning of Euler-Lagrange dynam- ics from data. *J. Comput. Appl. Math*, (2022), 114780. https://doi.org/10.1016/j.cam.2022.114780
11. C. Offen, Release of GitHub repository Christian-Offen/BEAConjugateSymplectic, Available from: https://github.com/Christian-Offen/BEAConjugateSymplectic.
12. C. Offen, S. Ober-Blöbaum, Symplectic integration of learned Hamiltonian systems. Poster presented at the workshop Theory of deep learning (9 August 2021 to 13 August 2021), Isaac Newton Institute, University of Cambridge, Cambridge, UK, 2021. Available from: https://github.com/Christian-Offen/symplectic-shadow-integration/raw/master/PosterWorkshopCambridge.pdf.
13. C. Offen, S. Ober-Blöbaum, Symplectic integration of learned Hamiltonian systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, **32** (2022). https://doi.org/10.1063/5.0065913
14. S. Sigurdsson, Linear multistep methods with variable matrix coefficients, *In Lecture Notes in Mathematics*, (1971), 327-331. Springer Berlin Heidelberg. https://doi.org/10.1007/BFb0069468
15. M. Vermeeren, Modified equations for variational integrators, *Numer. Math*, **137** (2017), 1001–1037. https://doi.org/10.1007/s00211-017-0896-4
16. E. T. Whittaker, S. W. McCrae, Hamiltonian systems and their integral invariants, *Cambridge Mathematical Library, Cambridge University Press,* https://doi.org/10.1017/CBO9780511608797.012