Multisymplectic Integration

A thesis presented in partial fulfilment
of the requirements for the degree of

Doctor of Philosophy
in
Mathematical Physics

at Massey University, Palmerston North,
New Zealand.

Brett Nicholas Ryland

2007

Abstract

Multisymplectic integration is a relatively new addition to the field of geometric integration, which is a modern approach to the numerical integration of systems of differential equations. Multisymplectic integration is carried out by numerical integrators known as multisymplectic integrators, which preserve a discrete analogue of a multisymplectic conservation law.

In recent years, it has been shown that various discretisations of a multi-Hamiltonian PDE satisfy a discrete analogue of a multisymplectic conservation law. In particular, discretisation in time and space by the popular symplectic Runge–Kutta methods has been shown to be multisymplectic. However, a multisymplectic integrator not only needs to satisfy a discrete multisymplectic conservation law, but it must also form a well-defined numerical method. One of the main questions considered in this thesis is that of when a multi-Hamiltonian PDE discretised by Runge–Kutta or partitioned Runge–Kutta methods gives rise to a well-defined multisymplectic integrator. In particular, multisymplectic integrators that are explicit are sought, since an integrator that is explicit will, in general, be well defined.

The first class of discretisation methods that I consider are the popular symplectic Runge–Kutta methods. These have previously been shown to satisfy a discrete analogue of the multisymplectic conservation law. However, these previous studies typically fail to consider whether or not the system of equations resulting from such a discretisation is well defined. By considering the semi-discretisation and the full discretisation of a multi-Hamiltonian PDE by such methods, I show the following:

- For Runge–Kutta (and for partitioned Runge–Kutta methods), the active variables in the spatial discretisation are the stage variables of the method, not the node variables (as is typical in the time integration of ODEs).

- The equations resulting from a semi-discretisation with periodic boundary conditions are only well defined when both the number of stages in the Runge–Kutta method and the number of cells in the spatial discretisation are odd. For other types of boundary conditions, these equations are not well defined in general.

- For a full discretisation, the numerical method appears to be well defined at first, but for some boundary conditions, the numerical method fails to accurately represent the PDE, while for other boundary conditions, the numerical method is highly implicit, ill-conditioned and impractical for all but the simplest of applications. An exception to this is the Preissman box scheme, whose simplicity avoids the difficulties of higher order methods.

- For a multisymplectic integrator, boundary conditions are treated differently in time and in space. This breaks the symmetry between time and space that is inherent in multisymplectic geometry.

The second class of discretisation methods that I consider are partitioned Runge–Kutta methods. Discretisation of a multi-Hamiltonian PDE by such methods has lead to the following two major results:

1. There is a simple set of conditions on the coefficients of a general partitioned Runge–Kutta method (which includes Runge–Kutta methods) such that a general multi-Hamiltonian PDE, discretised (either fully or partially) by such methods, satisfies a natural discrete analogue of the multisymplectic conservation law associated with that multi-Hamiltonian PDE.
2. I have defined a class of multi-Hamiltonian PDEs that, when discretised in space by a member of the Lobatto IIIA–IIIB class of partitioned Runge–Kutta methods, give rise to a system of explicit ODEs in time by means of a construction algorithm. These ODEs are well defined (since they are explicit), local, high order, multisymplectic and handle boundary conditions in a simple manner without the need for any extra requirements. Furthermore, by analysing the dispersion relation for these explicit ODEs, it is found that such spatial discretisations are stable.

From these explicit ODEs in time, well-defined multisymplectic integrators can be constructed by applying an explicit discretisation in time that satisfies a fully discrete analogue of the semi-discrete multisymplectic conservation law satisfied by the ODEs. Three examples of explicit multisymplectic integrators are given for the nonlinear Schrödinger equation, whereby the explicit ODEs in time are discretised by the 2-stage Lobatto IIIA–IIIB, linear–nonlinear splitting and real–imaginary–nonlinear splitting methods. These are all shown to satisfy discrete analogues of the multisymplectic conservation law, however, only the discrete multisymplectic conservation laws satisfied by the first and third multisymplectic integrators are local.

Since it is the stage variables that are active in a Runge–Kutta or partitioned Runge–Kutta discretisation in space of a multi-Hamiltonian PDE, the order of such a spatial discretisation is limited by the order of the stage variables. Moreover, the spatial discretisation contains an approximation of the spatial derivatives, and thus, the order of the spatial discretisation may be further limited by the order of this approximation. For the explicit ODEs resulting from an r-stage Lobatto IIIA–IIIB discretisation in space of an appropriate multi-Hamiltonian PDE, the order of this spatial discretisation is $r-1$ for $r \leq 10$; this is conjectured to hold for higher values of r. For $r=3$, I show that a modification to the initial conditions improves the order of this spatial discretisation. It is expected that a similar modification to the initial conditions will improve the order of such spatial discretisations for higher values of r.

Acknowledgements

First and foremost, I would like to say a big thank you to my main supervisor, Prof. Robert McLachlan, who has spent countless hours over the last decade (firstly as an undergraduate lecturer, then as my Master's supervisor and lastly as my Doctoral supervisor) enthusiastically instructing me in the theory and practice of geometric integration, numerical analysis, and various other fields of mathematics, all-the-while patiently explaining and re-explaining ideas that he had thought of since our previous meeting or getting overly excited whenever I brought a new result or interesting looking graph to his attention; hopefully I've taught you a few things along the way too.

Thanks also go to Robert's wife and two daughters for allowing Robert and me to have the occasional meeting while he was on parental leave, since, within a few weeks of becoming my supervisor (both for my Master's and my PhD), Robert announced that his wife was with child!

I would like to thank my parents for providing me with an education that has allowed me to reach this far, for their support and encouragement throughout my studies and for repeatedly moving and storing my stuff. A thank you also goes to my three brothers who have provided stiff competition for almost every challenge I have undertaken, encouraging me to strive even further than I might have otherwise.

My thanks go to my fellow PhD (or ex-PhD) friends, Priscilla, Dion, Philip, Patrick (Paddy) and Paul (PP) for numerous discussions over coffee and their entertaining observations of graduate life. A thank you also goes to my friends in the Phoenix medieval re-enactment company, Friday night role-playing group, canoe polo leagues and white water kayaking clubs, who have provided much in the way of entertainment and physical exercise when I have felt the need for procrastination.

Finally, I would also like to thank those who have been my friends since my early years at university. In particular, a special thank you goes to my dear friends, Miriam (who kindly provided some last minute grammar and spell checking), Nathan, Anita and Jane, who provide a relatively sane perspective on the world outside of university life and whose antics continue to keep me entertained.

Brett N. Ryland November 5, 2007.

Contents

List of Figures

Chapter 1

Introduction

Multisymplectic integration is a relatively new addition to the field of geometric integration, which is a modern approach to the numerical integration of systems of differential equations. Geometric integration is achieved through the use of numerical methods known as geometric integrators. The defining feature of geometric integrators is that they are derived from the system under consideration in such a way that the numerical solution exactly preserves one or more properties of the continuous solution. Some of the more desirable properties that geometric integrators may possess are time-reversibility, the preservation of symmetries, the conservation of first integrals, such as energy and momentum, and the conservation of various other properties, such as volume, symplecticity and multisymplecticity. The focus of this thesis is on multisymplectic integration by geometric integrators that conserve multisymplecticity, otherwise known as multisymplectic integrators.

It is generally accepted that the concepts of symplecticity and multisymplecticity date back to 1935 with the early, and independent, works of Hermann Weyl [76, 77] and Théophile De Donder [20] on the calculus of variations for multiple integrals. It was, however, Weyl who first used the term "symplectic" to describe systems with symplectic conservation laws [77]:

> The name "complex group" formerly advocated by me in allusion to line complexes, ... has become more and more embarrassing through collision with the word "complex" in the connotation of complex number. I therefore propose to replace it by the Greek adjective "symplectic."

For many years after the work of Weyl and De Donder, the concepts of symplecticity and multisymplecticity received only sporadic attention, which was generally in the context of Lagrangian field theories [21, 22, 28, 29, 46, 47, 74]. While numerical integrators continued to be developed during this period, these methods generally focussed on properties such as improved order, conservation of first integrals and simplicity of implementation; little attention was given to the concepts of symplecticity and multisymplecticity.

Symplecticity as a property of numerical integrators became of interest to the numerical integration community in the late 1980's, when it was discovered that many of the well-behaved methods that had been previously constructed, such as the implicit midpoint and Störmer–Verlet methods, conserved symplecticity. In particular, the conditions for the popular Runge–Kutta methods (which were first constructed around the turn of the twentieth century and later characterised by a tableau of coefficients [16]) to conserve symplecticity were discovered independently in 1988 [48, 67] and 1989 [73]. Nearly a decade later, the ideas that were being used to construct symplectic integrators for Hamiltonian ODEs began to be applied to Hamiltonian PDEs. This has led to a dramatic surge in the amount of interest in the developing field of multisymplectic integration, as is evidenced by the number of recent publications in this area [2, 3, 5, 6, 7, 8, 10, 11, 13, 14, 15, 18, 27, 30, 37, 39, 40, 41, 42, 43, 45, 49, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 62, 63, 64, 65, 66, 71, 75].

Currently, there are two approaches to the field of multisymplectic integration. The first approach is the traditional Lagrangian approach, which, in the context of multisymplecticity, dates back to the work of Weyl and De Donder. This approach is couched in the language of differential geometry, which is an abstract coordinate free description of calculus on manifolds (an introduction to this approach is given in Section 1.3). The Lagrangian approach is based upon the idea that variational principles underlie the fundamental laws of force balance, i.e., $\boldsymbol{F} = m\boldsymbol{a}$. These variational principles allow the equations of motion of a system to be derived from the Lagrangian, which is a function describing the dynamics of that system. Furthermore, these variational principles have close ties with Noether theory and, as such, symmetries in the system appear as conservation laws.

To form a multisymplectic integrator from such a variational description, a discrete approximation of a Lagrangian PDE is considered. Along with this so-called discrete Lagrangian, a discrete analogue of the apparatus used in the theory of continuous Lagrangians is constructed. In particular, the discrete analogue of Hamilton's variational principle gives rise to algorithms known as discrete variational integrators, which conserve multisymplecticity (see Section 1.4.5).

The second approach is based on an extension of the theory of Hamiltonian ODEs to Hamiltonian PDEs, whereby the PDE is written in a so-called multi-Hamiltonian form (see Section 1.2), which was introduced in 1997 by Thomas Bridges [7, 8]. We refer to a Hamiltonian PDE written this way as a multi-Hamiltonian PDE. This approach is particularly suited to the study of wave-like PDEs and has led to a deeper understanding of the properties of wave propagation [8].

Multisymplectic integrators based on this approach are constructed by discretising a multi-Hamiltonian PDE by methods that preserve the multisymplectic structure of the PDE. Shortly after the introduction of this multi-Hamiltonian form it was shown that discretisation by symplectic Runge–Kutta methods preserves the multisymplectic structure of a multi-Hamiltonian PDE [63]. In particular, the Preissman box scheme, which was introduced by Preissman in 1960 and is used extensively in fluid dynamics applications (such as weather prediction), has been shown to be multisymplectic [13]. Since then, various other discretisation methods have been shown to preserve the multisymplectic structure of multi-Hamiltonian PDEs [38, 66]. It is this approach that will be considered throughout most of this thesis.

Since multisymplectic integrators can be constructed via each of these approaches, a question that naturally arises is whether one approach is more fundamental than the other. As yet, the answer to this question has not been decided. On the one hand, the Lagrangian approach may be considered to be more fundamental since it is based on variational principles, and may be generalised to general relativity more easily than the Hamiltonian approach. On the other hand, the Hamiltonian approach may be considered more fundamental since it is based on the concept of energy, which has close ties with quantum mechanics. Fortunately, in many cases, these two approaches are equivalent and the question is moot.

1.1 General overview of this thesis

The remainder of this chapter is laid out as follows. In Section 1.2, I give an introduction to the Hamiltonian approach to the theory of multisymplectic integration, including a description of the various types of conservation law. In particular, the definition of a multi-Hamiltonian PDE and its corresponding multisymplectic conservation law are given. In Section 1.3, I give a description of the Lagrangian approach to the theory of multisymplectic integration, including an introduction to the apparatus of differential geometry that is required for this approach. Following this, I give an overview of several commonly used symplectic and multisymplectic integrators in Section 1.4. Lastly, in Section 1.5, I give an overview of the advantages that multisymplectic integrators enjoy over other types of numerical integrators and list a selection of PDEs that have had multisymplectic integrators constructed for them.

In Chapter 2, I discuss the discretisation of a multi-Hamiltonian PDE by symplectic Runge–Kutta methods using the nonlinear wave equation as the primary example. After showing that the discrete set of equations satisfies a discrete analogue of the multisymplectic conservation law in Section 2.1, I proceed to determine the conditions when such discretisations form well-defined numerical methods. Firstly, in Section 2.2, spatial semi-discretisations are considered and it is shown that for periodic boundary conditions, a necessary condition for the discretisation to be well defined is that both the number of grid points and the number of stages in the Runge–Kutta method must be odd. Furthermore, for other boundary conditions the discretisation fails to be well defined in general. Secondly, in Section 2.3, full discretisation in time and space is considered and found to be well defined if boundary conditions are chosen appropriately. However, for some choices of boundary conditions the numerical integrator does not represent the PDE, while for other choices of boundary conditions the numerical integrator is impractical in general.

In Chapter 3, partitioned Runge–Kutta discretisations of a multi-Hamiltonian PDE are considered. I begin this chapter by giving the general definition of a partitioned Runge–Kutta method. Following this, I define a simple set of conditions on the coefficients of these methods such that the discretisation (or semi-discretisation) of a multi-Hamiltonian PDE by such methods satisfies a discrete analogue of the multisymplectic conservation law. Then, in Section 3.4, I define a class of multi-Hamiltonian PDEs such that a discretisation in space by the Lobatto IIIA–IIIB class of partitioned Runge–Kutta methods leads to a system of explicit ODEs it time. This is shown by way of a construction algorithm. Furthermore, I show that these ODEs handle boundary conditions in a local and simple manner. In order to demonstrate the necessity of the conditions that define this class of multi-Hamiltonian PDEs, several examples of multi-Hamiltonian PDEs that lie in this class are given along with several examples that do not.

In Chapter 4, I describe how explicit multisymplectic integrators may be constructed from the explicit ODEs in Section 3.4 by an appropriate discretisation in time. The primary example used in this chapter is the nonlinear Schrödinger (NLS) equation. The first discretisation method considered is a Lobatto IIIA–IIIB discretisation in time with coefficients satisfying the conditions for the conservation of multisymplecticity. The second and third discretisation methods considered are splitting methods. It is shown that the first and third multisymplectic integrators satisfy local discrete analogues of the multisymplectic conservation law, whereas the second multisymplectic integrator satisfies a non-local discrete analogue of the multisymplectic conservation law. This is followed by a discussion on the ability of these three multisymplectic integrators to preserve the three basic conservation laws (energy, momentum and norm) that are possessed by the NLS equation.

In Chapter 5, I give an analysis of the dispersion relation and the order of Runge–Kutta and partitioned Runge–Kutta discretisations of a multi-Hamiltonian PDE. In particular, in Section 5.1, I give the dispersion relation for a 3-stage Lobatto IIIA–IIIB discretisation in space of the nonlinear wave equation. This dispersion relation exhibits more solutions than the dispersion relation of the continuous PDE, which is typically an indication of an unstable mode in the solution and limits the step size that may be taken by the integrator. However, for Runge–Kutta and partitioned Runge–Kutta methods, it is shown that these extra solutions do not correspond to unstable modes and that these methods are stable. In Section 5.2, the order of the Lobatto IIIA–IIIB discretisation in space of a multi-Hamiltonian PDE is considered. It is shown that this discretisation (and the Gaussian Runge–Kutta discretisations that have previously been considered) only has the order of the stage variables in the method. However, by a careful selection of the initial conditions, this order may be increased.

Lastly, in Chapter 6, I make my concluding remarks, followed by a list of open questions arising from the results of this thesis.

1.2 Hamiltonian ODEs and PDEs

The Hamiltonian approach to the field of multisymplectic integration is based upon the concept of energy. That is, for an ODE, a function (the *Hamiltonian*) representing the total energy of the system is constructed, from which the equations of motion of the system can be derived.

The dependent variables $\boldsymbol{q} = \{q^1, \ldots, q^n\}$ (referred to as generalised coordinates) and $\boldsymbol{p} = \{p_1, \ldots, p_n\}$ (referred to as generalised conjugate momenta) are introduced as functions of an independent time-like variable, t. Together, the generalised coordinates and conjugate momenta define the phase space $(\boldsymbol{q}, \boldsymbol{p}) \in \mathbb{R}^{2n}$. On this phase space, the Hamiltonian is defined as the function

$$H = H(\boldsymbol{q}, \boldsymbol{p}, t). \tag{1.1}$$

The equations of motion of a Hamiltonian system are given by *Hamilton's equations*,

$$\frac{\mathrm{d}q^i}{\mathrm{d}t} = \frac{\partial H}{\partial p_i}, \qquad \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial H}{\partial q^i}, \tag{1.2}$$

which describe the evolution in time of a point in phase space.

Eq. (1.1) is the general form of the Hamiltonian, but in most systems of interest the Hamiltonian is autonomous, that is, it does not depend explicitly on t, and, as a consequence, the energy of the system is constant in time. In many situations, the Hamiltonian is also separable, meaning that it can be written as the sum of functions on mutually exclusive sets of variables. For example, in particle mechanics the Hamiltonian can be written as $H(\boldsymbol{q}, \boldsymbol{p}) = T(\boldsymbol{p}) + V(\boldsymbol{q})$, where $T(\boldsymbol{p})$ and $V(\boldsymbol{q})$ are kinetic and potential energy, respectively.

It can be immediately noted from Hamilton's equations that autonomous Hamiltonian systems have the property,

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \sum_{i=1}^{n} \left(\frac{\partial H}{\partial q^i}\frac{\mathrm{d}q^i}{\mathrm{d}t} + \frac{\partial H}{\partial p_i}\frac{\mathrm{d}p_i}{\mathrm{d}t} \right) \equiv 0, \tag{1.3}$$

which corresponds to the conservation of energy by solutions of the autonomous Hamiltonian system. Unless otherwise stated, the Hamiltonian systems considered throughout the remainder of this thesis may be assumed to be autonomous.

Hamilton's equations can be written more concisely by writing the coordinates and momenta as a single vector $\boldsymbol{z} = (q^1, \ldots, q^n, p_1, \ldots, p_n)$, then a point in phase space evolves in time according to

$$\boldsymbol{z}_t = \boldsymbol{J}^{-1}\nabla_{\boldsymbol{z}} H(\boldsymbol{z}), \tag{1.4}$$

where

$$\boldsymbol{J}^{-1} = \begin{bmatrix} \boldsymbol{0}_n & \boldsymbol{I}_n \\ -\boldsymbol{I}_n & \boldsymbol{0}_n \end{bmatrix} \tag{1.5}$$

and $\boldsymbol{z}_t$ is the time-derivative of $\boldsymbol{z}$. $\boldsymbol{J}^{-1}$ is often referred to as the symplectic matrix since the Lie algebra of the symplectic group of degree $2n$ over a field consists of the matrices $\boldsymbol{A}$ (with elements in the field) satisfying $\boldsymbol{J}^{-1}\boldsymbol{A} + \boldsymbol{A}^T\boldsymbol{J}^{-1} = 0$.

Hamiltonian PDEs are the extension of Hamiltonian ODEs to the situation where the dependent variables, $\boldsymbol{z}$, are functions of multiple variables (typically 1 time-like variable, t, and 1 space-like variable, x). In these PDEs, each of the dimensions are treated on an equal footing.

For a Hamiltonian PDE, the total energy in the system is no longer represented by a single function. Instead we have two functions, $E(\boldsymbol{z})$ and $F(\boldsymbol{z})$, which describe the local energy density and local energy flux in the system. Together, the energy density and energy flux describe how the distribution of energy in a PDE varies over time. However, while the amount of energy in an autonomous Hamiltonian PDE will vary locally in general, the total amount of energy in the system will remain constant.

The natural generalisation of Hamilton's equations for a Hamiltonian PDE is to write the PDE as a system of first order equations in the so-called *multi-Hamiltonian form* [7, 8]

$$\boldsymbol{K}\boldsymbol{z}_t + \boldsymbol{L}\boldsymbol{z}_x = \nabla_{\boldsymbol{z}} S(\boldsymbol{z}), \tag{1.6}$$

where $\boldsymbol{z} \in \mathbb{R}^n$, $\boldsymbol{K}$ and $\boldsymbol{L}$ are skew-symmetric matrices in $\mathbb{R}^{n\times n}$ and $S(\boldsymbol{z})$ is a smooth function. Note that $S(\boldsymbol{z})$ is autonomous here (it has no explicit t or x dependence), Hamiltonian PDEs written in the form of Eq. (1.6) but with a non-autonomous $S(\boldsymbol{z})$ have not been considered in the literature and will not be considered here either. A Hamiltonian PDE written in the form of Eq. (1.6) is referred to as a *multi-Hamiltonian PDE.*

As an example, the nonlinear wave equation, $u_{tt} = u_{xx} - V'(u)$, can be written as a multi-Hamiltonian PDE with

$$\boldsymbol{z} = \begin{bmatrix} u \\ v \\ w \end{bmatrix}, \quad \boldsymbol{K} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{L} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \tag{1.7}$$

and $S(\boldsymbol{z}) = -V(u) + \frac{1}{2}(w^2 - v^2)$. This example has been used extensively in the literature and is one of a number of multi-Hamiltonian PDEs that are considered in this thesis. While the multi-Hamiltonian form of many important PDEs are known, there is currently no systematic method of constructing the multi-Hamiltonian form of a given Hamiltonian PDE and it is not known whether this is even possible in general.

An important concept in the study of multisymplectic integration is that of the first variation of a PDE. For a multi-Hamiltonian PDE, this is given by

$$\boldsymbol{K}\mathrm{d}\boldsymbol{z}_t + \boldsymbol{L}\mathrm{d}\boldsymbol{z}_x = \boldsymbol{D}_{\boldsymbol{zz}}S(\boldsymbol{z})\mathrm{d}\boldsymbol{z}, \tag{1.8}$$

where $\boldsymbol{K}$ and $\boldsymbol{L}$ are the same skew-symmetric matrices as above and $\boldsymbol{D}_{\boldsymbol{zz}}S(\boldsymbol{z})$ is the Hessian of $S(\boldsymbol{z})$.

In higher dimensions, the concept of a multi-Hamiltonian PDE is extended further to include PDEs of the form

$$\sum_{\alpha=1}^{N} \boldsymbol{K}^{\alpha}\boldsymbol{z}_{,x_\alpha} = \nabla_{\boldsymbol{z}}S(\boldsymbol{z}), \tag{1.9}$$

where $\boldsymbol{z}_{,x_\alpha}$ indicates the partial derivative of $\boldsymbol{z}$ with respect to the variable x_α. Eq. (1.9) reduces to Eq. (1.6) when $x_\alpha = (t, x)$. The first variation of Eq. (1.9) is given by

$$\sum_{\alpha=1}^{N} \boldsymbol{K}^{\alpha}\mathrm{d}\boldsymbol{z}_{,x_\alpha} = \boldsymbol{D}_{\boldsymbol{zz}}S(\boldsymbol{z})\mathrm{d}\boldsymbol{z}. \tag{1.10}$$

The focus of this manuscript is, for the main part, restricted to the consideration of multi-Hamiltonian PDEs of the form of Eq. (1.6) and their associated multisymplectic conservation laws.

1.2.1 Conservation laws

Conservation laws can be classified into four distinct categories depending on the whether the system studied is an ODE or a PDE and whether the conservation law is evaluated on solutions to the system or on solutions to the first variation of the system.

	ODE	PDE
Conservation Law	$I_t = 0$	$E_t + F_x = 0$
Differential Conservation Law	$\omega_t = 0$	$\omega_t + \kappa_x = 0$

Furthermore, these conservation laws, which are described in the sections below, are continuous and apply to the ODE or PDE under consideration; a numerical integrator is said to preserve a given property when it satisfies a discrete analogue of the corresponding continuous conservation law.

ODE conservation laws

ODE conservation laws can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}I(u(t)) = 0, \tag{1.11}$$

where I is the conserved quantity and $u(t)$ is a solution of the ODE.

In most applications where the ODE is derived from a physical system, the conserved quantity is energy and the integral is the Hamiltonian. Integrators for such systems are known (not surprisingly) as energy-preserving integrators.

An example of a numerical method that preserves a discrete analogue of this type of conservation law is the implicit midpoint method for a rigid body. The kinetic energy of a rigid body is given by

$$H(p_1, p_2, p_3) = \frac{1}{2}\left(\frac{p_1^2}{I_1} + \frac{p_2^2}{I_2} + \frac{p_3^2}{I_3}\right), \tag{1.12}$$

where p_i represent the angular momentum in the body frame and I_i are the principal moments of inertia. The kinetic energy is conserved since the implicit midpoint method belongs to the Gaussian Runge–Kutta class of methods, which conserve all quadratic invariants [33].

PDE conservation laws

A PDE conservation law consists of two functions evaluated on the solutions, $u(x,t)$, of the PDE, which together describe a locally conserved quantity, i.e.,

$$\frac{\partial}{\partial t}E(u(t,x)) + \frac{\partial}{\partial x}F(u(t,x)) = 0, \tag{1.13}$$

where the functions $E(u(t,x))$ and $F(u(t,x))$ describe the local density and flux of the conserved quantity, respectively.

As mentioned in Section 1.2, a multi-Hamiltonian PDE (1.6) possesses an energy conservation law of this type. However, due to the nature of Eq. (1.6), such a PDE also possesses a second conservation law of this type where the conserved quantity is momentum. These conservation laws are defined as follows:

- Energy:

$$\partial_t E(\boldsymbol{z}) + \partial_x F(\boldsymbol{z}) = 0, \tag{1.14}$$

  where $E(\boldsymbol{z}) = S(\boldsymbol{z}) + (\partial_x \boldsymbol{z})^T \boldsymbol{L}\boldsymbol{z}$ is the energy density and $F(\boldsymbol{z}) = -(\partial_t \boldsymbol{z})^T \boldsymbol{L}\boldsymbol{z}$ is the energy flux.

- Momentum:

$$\partial_t I(\boldsymbol{z}) + \partial_x G(\boldsymbol{z}) = 0, \tag{1.15}$$

  where $I(\boldsymbol{z}) = -(\partial_x \boldsymbol{z})^T \boldsymbol{K}\boldsymbol{z}$ is the momentum density and $G(\boldsymbol{z}) = S(\boldsymbol{z}) + (\partial_t \boldsymbol{z})^T \boldsymbol{K}\boldsymbol{z}$ is the momentum flux.

A property of PDE conservation laws is that they can be integrated over space (with suitable boundary conditions) to give an ODE conservation law,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int E(u(t,x))\mathrm{d}x\right) = 0, \tag{1.16}$$

where the globally conserved quantity is the integral $\int E(u(t,x))\mathrm{d}x$.

A class of numerical methods that preserve a discrete analogue of a PDE conservation law are those known as finite volume methods, where the space-time manifold is broken into small volume elements and fluxes occur at the boundaries of each element in such a way that the conservation law is satisfied by each element.

It should be noted, however, that PDE conservation laws are not a property exclusive to multi-Hamiltonian PDEs, e.g., Burger's equation: $u_t + (\frac{1}{2}u^2 - u_x)_x = 0$ possesses a mass conservation law.

ODE differential conservation laws

In contrast to the above conservation laws, differential conservation laws are not evaluated on solutions of the system. Instead, along such solutions of the system, the differential conservation law is evaluated on functions satisfying the first variation of the system.

Along solutions $u(t)$ of an ODE, a differential conservation law can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\omega(\mathrm{d}u(t)) = 0, \tag{1.17}$$

where the 2-form ω is given by $\omega = \boldsymbol{A}(u)\mathrm{d}u \wedge \mathrm{d}u$, for some matrix $\boldsymbol{A}(u)$, and $\mathrm{d}u$ satisfies the first variation of the ODE. This is illustrated in figure 1.1.

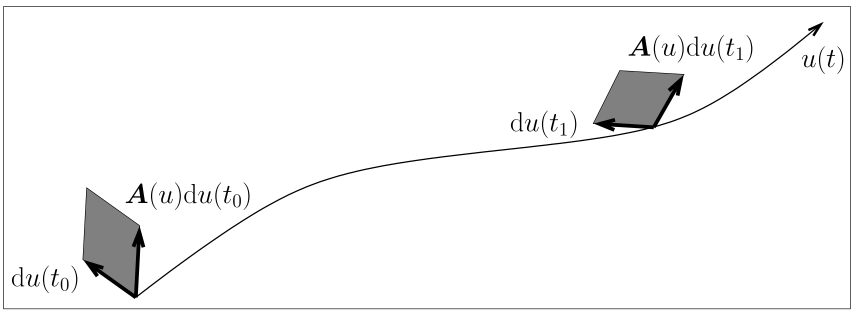


Figure 1.1: Differential conservation laws are evaluated on differentials, $\mathrm{d}u$, along a solution of the ODE. The differentials satisfy the first variation of the ODE.

If the ODE is a Hamiltonian ODE, then $\boldsymbol{A}(u)$ is the constant skew-symmetric matrix $\frac{1}{2}\boldsymbol{J}$, where $\boldsymbol{J}$ is defined by Eq. (1.5). The 2-form ω is then given by

$$\omega = \frac{1}{2}\boldsymbol{J}\mathrm{d}\boldsymbol{z} \wedge \mathrm{d}\boldsymbol{z}, \tag{1.18}$$

and Eq. (1.17) is called a *symplectic conservation law*.

The leapfrog integrator (see Section 1.4.3) is an example of a numerical method that conserves symplecticity, i.e., the value of ω is preserved when the variables in the system are updated from one time level to the next.

PDE differential conservation laws

In a similar manner to PDE conservations laws, the PDE version of a differential conservation law corresponds to the local conservation of some quantity.

Along solutions $u(t,x)$ of a PDE, a differential conservation law is given by

$$\frac{\partial}{\partial t}\omega(\mathrm{d}u(t,x)) + \frac{\partial}{\partial x}\kappa(\mathrm{d}u(t,x)) = 0, \tag{1.19}$$

where the 2-forms ω and κ are given by $\omega = \boldsymbol{A}(u)\mathrm{d}u \wedge \mathrm{d}u$ and $\kappa = \boldsymbol{B}(u)\mathrm{d}u \wedge \mathrm{d}u$ for some matrices $\boldsymbol{A}(u)$ and $\boldsymbol{B}(u)$ and $\mathrm{d}u$ satisfies the first variation of the PDE.

If the PDE is a multi-Hamiltonian PDE (1.6) then $\boldsymbol{A}(u)$ and $\boldsymbol{B}(u)$ are the constant skew-symmetric matrices $\frac{1}{2}\boldsymbol{K}$ and $\frac{1}{2}\boldsymbol{L}$, respectively. The 2-forms ω and κ are then given by

$$\begin{aligned}\omega &= \frac{1}{2}\boldsymbol{K}\mathrm{d}\boldsymbol{z} \wedge \mathrm{d}\boldsymbol{z},\\ \kappa &= \frac{1}{2}\boldsymbol{L}\mathrm{d}\boldsymbol{z} \wedge \mathrm{d}\boldsymbol{z},\end{aligned} \tag{1.20}$$

and Eq. (1.19) is referred to as a *multisymplectic conservation law*.

Unlike symplectic conservation laws, which describe the conservation of the total amount of symplecticity in the system, multisymplectic conservation laws describe the conservation of the amount of multisymplecticity at a point in time and space, i.e., at each point in time and space, the change in the amount of multisymplecticity in time is balanced by the change in the amount of multisymplecticity in space.

Multisymplectic conservation laws are often mistakenly identified in the literature as locally conserving symplecticity [13, 64]. The reason for this is that integration of a multisymplectic conservation law over space, with suitable boundary conditions, gives a symplectic conservation law in time, i.e.,

$$\int (\omega_t + \kappa_x)\mathrm{d}x = \partial_t \int \omega \mathrm{d}x = 0, \tag{1.21}$$

where the symplectic 2-form is the integral $\int \omega \mathrm{d}x$. Similarly, integration of a multisymplectic conservation law over time, with suitable boundary conditions, gives a differential conservation law in space for the 2-form $\int \kappa \mathrm{d}t$. However, this pair of ODE differential conservation laws (one in time and one in space) does not necessarily imply the conservation of multisymplecticity as each of the ODE differential conservation laws are global (the first in space and the second in time), whereas the conservation of multisymplecticity is a property that is local in both time and space.

The Preissman box scheme is an example of a numerical integrator that preserves a discrete analogue of a multisymplectic conservation law (a so-called *discrete multisymplectic conservation law*) and is the result of applying finite differences in both time and space. Applying the Preissman box scheme to Eq. (1.6) on a grid with constant spacing gives

$$\frac{1}{2\Delta x}\boldsymbol{K}(\boldsymbol{z}_{i+1}^{n+1}+\boldsymbol{z}_{i+1}^{n}-\boldsymbol{z}_{i}^{n+1}-\boldsymbol{z}_{i}^{n})+\frac{1}{2\Delta t}\boldsymbol{L}(\boldsymbol{z}_{i+1}^{n+1}+\boldsymbol{z}_{i}^{n+1}-\boldsymbol{z}_{i+1}^{n}-\boldsymbol{z}_{i}^{n}) \\ =\nabla_{\boldsymbol{z}}S(\frac{1}{4}(\boldsymbol{z}_{i+1}^{n+1}+\boldsymbol{z}_{i}^{n+1}+\boldsymbol{z}_{i+1}^{n}+\boldsymbol{z}_{i}^{n})), \quad (1.22)$$

where $\boldsymbol{z}_i^n$ represents the variables at the node at $i\Delta x$ in space and $n\Delta t$ in time.

In higher dimensions a multi-Hamiltonian PDE written in the form of Eq. (1.9) possesses a multisymplectic conservation law given by

$$\sum_{\alpha=1}^{N}\omega^{\alpha}_{,x_\alpha}=0, \quad (1.23)$$

where $\omega^\alpha=\frac{1}{2}\boldsymbol{K}^\alpha\mathrm{d}\boldsymbol{z}\wedge\mathrm{d}\boldsymbol{z}$.

1.3 Lagrangian mechanics

The Lagrangian approach to the field of multisymplectic integration is based upon the idea that variational principles underlie the fundamental laws of force balance. That is, a function (the *Lagrangian*) is constructed which describes the dynamics of the system. The equations of motion of this system may then be derived by considering variations of the solution that hold the integral of the Lagrangian stationary. This idea is spelled out in more detail below.

In general, the motion described by an ODE can be formulated on an n-dimensional configuration manifold Q (with coordinates $\boldsymbol{q}=(q^1,\ \ldots,\ q^n)$) together with a $2n$-dimensional tangent bundle TQ (with coordinates $(\boldsymbol{q},\dot{\boldsymbol{q}})$, where $\dot{\boldsymbol{q}}$ indicates the derivative of $\boldsymbol{q}$ with respect to time). On this tangent bundle, the Lagrangian, $L(\boldsymbol{q},\dot{\boldsymbol{q}},t)$, which is often of the form kinetic energy minus potential energy, is constructed. Then the action functional is defined as the definite integral of the Lagrangian, i.e.,

$$S(\boldsymbol{q}(t))=\int_a^b L(\boldsymbol{q},\dot{\boldsymbol{q}},t)\mathrm{d}t. \quad (1.24)$$

Now, the *variational principle of Hamilton* states that taking variations, $\delta\boldsymbol{q}$, of the solution, $\boldsymbol{q}(t)$, with the endpoints fixed, we can find solutions of the ODE which leave the action functional stationary. That is, taking variations of the solution gives

$$\mathrm{d}S(\boldsymbol{q}(t)) \cdot \delta\boldsymbol{q}(t) = \int_a^b \delta q^i \left(\frac{\partial L}{\partial q^i} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}^i} \right) \mathrm{d}t + \frac{\partial L}{\partial \dot{q}^i}\delta q^i \bigg|_a^b, \tag{1.25}$$

and since the only constraint on δq^i is that the endpoints are fixed, the last term disappears, which gives the *Euler–Lagrange equations*,

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = 0. \tag{1.26}$$

Associated with the configuration bundle Q is the cotangent bundle T^*Q with coordinates $(\boldsymbol{q}, \boldsymbol{p}) = (q^1, \ \ldots, \ q^n, \ p_1, \ \ldots, \ p_n)$, which is the phase space constructed in Section 1.2. Under suitable non-degeneracy conditions on L, the Legendre transform defined by

$$p_i = (\partial L / \partial \dot{q}^i) \tag{1.27}$$

is a diffeomorphism from TQ to T^*Q, which maps the Euler–Lagrange equations to Hamilton's equations. This allows one to switch between Hamiltonian and Lagrangian settings via the equation

$$H(\boldsymbol{q}, \boldsymbol{p}, t) = \sum_{i=1}^{n} p_i \dot{q}^i - L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t). \tag{1.28}$$

It is the Legendre transform that typically gives physical meaning to the conjugate momenta.

1.3.1 Differential geometry for ODEs

A more concise way of working with Lagrangian and Hamiltonian mechanics is to use the coordinate free language of differential geometry. A good introduction to differential geometry is given in [55], with more advanced topics in [30, 52].

In Hamiltonian mechanics, the Hamiltonian is defined on the cotangent bundle, T^*Q. If the configuration bundle, Q, is a manifold, then T^*Q together with the symplectic 2-form $\omega = \mathrm{d}q^i \wedge \mathrm{d}p_i$ define a symplectic manifold (T^*Q, ω).

We say that a vector field, X, on T^*Q is Hamiltonian if there is a function, $H : T^*Q \mapsto \mathbb{R}$, such that

$$\imath_X \omega = \mathrm{d}H, \tag{1.29}$$

where the notation $\imath_X \alpha$ indicates the interior product of the vector field X and the k-form α; it is defined by

$$\imath_X \alpha(v_2, \ldots, v_k) = \alpha(X, v_2, \ldots, x_k). \tag{1.30}$$

The Hamiltonian vector field is then written as X_H, and Hamilton's equations for the evolution of the system are given by

$$\dot{\boldsymbol{z}} = X_H(\boldsymbol{z}). \tag{1.31}$$

Now, given two symplectic manifolds (P_1, ω_1) and (P_2, ω_2), a symplectic transformation (or map) is defined to be a C^∞-mapping, $\phi : P_1 \mapsto P_2$, such that $\phi^*\omega_2 = \omega_1$. That is, for each $z \in P_1$ and for all $v, w \in T_z P_1$, the identity

$$\omega_{1z}(v, w) = \omega_{2\phi(z)}(T_z\phi \cdot v, T_z\phi \cdot w) \tag{1.32}$$

holds. Thus the flow, $\varphi_t(\boldsymbol{z})$, of the Hamiltonian vector field X_H above, consists of symplectic transformations from the symplectic manifold to itself,

$$\varphi_t^*\omega = \omega. \tag{1.33}$$

This can be seen from

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t^*\omega = \varphi_t^*\pounds_X\omega = 0, \tag{1.34}$$

where $\pounds_X\omega = \imath_X\mathrm{d}\omega + \mathrm{d}\imath_X\omega = \mathrm{d}\imath_X\omega$ and $\imath_X\omega$ is closed ($\mathrm{d}\imath_X\omega = \mathrm{dd}H = 0$). Hence we have the result that symplecticity is conserved by the flow. A consequence of this is that energy is also conserved, i.e., $H \circ \varphi_t = H$ (where H is defined). This can be seen by noting that

$$\dot{H} = \imath_X\mathrm{d}H = \imath_X\imath_X\omega = 0, \tag{1.35}$$

since a double interior product over the same vector field is identically zero.

In Lagrangian mechanics, we have a Lagrangian defined on the tangent bundle, TQ. To show that symplecticity is conserved, a *Lagrangian vector field* that preserves a symplectic 2-form is sought. The standard method of obtaining a Lagrangian vector field is to define a canonical 1-form on the cotangent bundle and pull it back to the tangent bundle by a Legendre transform.

That is, the canonical 1-form θ_0 on T^*Q is defined by

$$\theta_0(\alpha_q)w_{\alpha_q} \equiv \alpha_q \cdot T\pi_Q w_{\alpha_q}, \tag{1.36}$$

where $\alpha_q \in T_q^*Q$, $w_{\alpha_q} \in T_{\alpha_q}T^*Q$ and $\pi_Q : T^*Q \mapsto Q$ is the canonical projection. Then the Legendre transformation, $\mathbb{F}L : TQ \mapsto T^*Q$, is defined by differentiation of the Lagrangian, i.e.,

$$\mathbb{F}L(v_q)w_q \equiv \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} L(v_q + \epsilon w_q). \tag{1.37}$$

The Lagrange 1-form can now be defined on TQ as

$$\theta_L \equiv \mathbb{F}L^*\theta_0, \tag{1.38}$$

which, in turn, defines the Lagrange 2-form as

$$\omega_L = -\mathrm{d}\theta_L. \tag{1.39}$$

If the Legendre transform is a local diffeomorphism, then ω_L is a symplectic 2-form and there exists a unique vector field (called the *Lagrange vector field* on TQ) whose flow, F_t, preserves ω_L, i.e.,

$$F_t^*\omega_L = \omega_L. \tag{1.40}$$

It is claimed, in [52], that it is advantageous to stay completely on the Lagrangian side by obtaining the Lagrange 1-form on TQ directly from the variational principle. By removing the constraint on Eq. (1.25) that the end points are fixed, it is seen that the term $(\partial L/\partial \dot{q}^i)\mathrm{d}q^i$ is the Lagrange 1-form on TQ. This is formalised in Theorem 2.1 of [52], which I now quote.

Theorem 2.1 *Given a C^k Lagrangian L, $k \geq 2$, there exists a unique C^{k-2} mapping $D_{EL}L : \ddot{Q} \to T^*Q$, defined on the second order submanifold*

$$\ddot{Q} \equiv \left\{ \frac{\mathrm{d}^2 q}{\mathrm{d}t^2}(0) \middle| \; q \text{ a } C^2 \text{ curve in } Q \right\}$$

of TTQ, and a unique C^{k-1} 1-form θ_L on TQ, such that, for all C^2 variations $q_\epsilon(t)$,

$$\mathrm{d}S(q(t)) \cdot \delta q(t) = \int_a^b D_{EL}L\left(\frac{\mathrm{d}^2 q}{\mathrm{d}t^2}\right) \cdot \delta q \mathrm{d}t + \theta_L\left(\frac{\mathrm{d}q}{\mathrm{d}t}\right) \cdot \hat{\delta q}\Big|_a^b,$$

where

$$\delta q(t) \equiv \frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=0} q_\epsilon(t), \qquad \hat{\delta q}(t) \equiv \frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=0} \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} q_\epsilon(t).$$

*The 1-form so defined is called the **Lagrange 1-form**.*

As before, the symplectic 2-form is then just the Lagrange 2-form obtained by taking the exterior derivative of the Lagrange 1-form, i.e.,

$$\omega_L \equiv -\mathrm{d}\theta_L. \tag{1.41}$$

The variational principle, when applied to a Lagrangian that is regular, gives coordinate independent second order ODEs, from which a vector field on TQ is generated. By considering the flow, F_t, of this vector field, one can find that the value of the action functional at time t, when restricted to the space of solutions of the variational principle, is given by

$$S_t = \int_0^t L(\boldsymbol{q}(s), \dot{\boldsymbol{q}}(s))\mathrm{d}s, \tag{1.42}$$

where $(\boldsymbol{q}(s), \dot{\boldsymbol{q}}(s)) = F_s(v_q)$ and $v_q \in TQ$.

Applying this to Theorem 2.1 of [52] gives the equation

$$\mathrm{d}S_t = F_t^* \theta_L - \theta_L. \tag{1.43}$$

Taking an exterior derivative of the above equation gives

$$\mathrm{d}\mathrm{d}S_t = -F_t^* \omega_L + \omega_L \equiv 0, \tag{1.44}$$

thus we again find that

$$F_t^* \omega_L = \omega_L, \tag{1.45}$$

that is, the solutions of the variational principle evolve in a manner that conserves symplecticity.

1.3.2 Differential geometry for PDEs

The introduction to differential geometry given in the previous section is sufficient for Lagrangian ODEs. However, to work with Lagrangian PDEs we must extend our notation, in particular we need the concept of jets and jet bundles.

Let X be an oriented manifold of dimension $n + 1$ with coordinates x^μ, for $\mu = 0, 1, \ldots, n$, corresponding to time plus n spatial dimensions. Let $\pi_{XY} : Y \mapsto X$ be a fibre bundle over X, where Y has fibre dimension N and fibre coordinates y^A, for $A = 1, \ldots, N$. Y is the multisymplectic equivalent of the configuration bundle Q in symplectic geometry.

Next, we can define the first jet bundle $J^1(Y)$ over Y to be the affine bundle over Y whose fibres over $y \in Y_x := \pi_{XY}^{-1}(x)$ consist of the linear mappings $\gamma : T_x X \mapsto T_y Y$ such that

$$T\pi_{XY} \circ \gamma = \text{Identity on } T_x X. \tag{1.46}$$

With the above coordinates on X and Y, $J^1(Y)$ has fibre dimension $n \times N$ with coordinates ${v^A}_\mu$ and the jet bundle is the multisymplectic equivalent of the tangent bundle, TQ.

Now, if $\phi : X \mapsto Y$ is a section of π_{XY}, then its tangent map $T_x\phi$ is an element of $J^1(Y)_{\phi(x)}$ and the map $x \mapsto T_x\phi$ defines a section of $J^1(Y)$ regarded as a bundle over X. This section, written as $j^1(\phi)$, is called the first jet of ϕ and is expressed in coordinates as

$$j^1(\phi) : x^\mu \mapsto (x^\mu, \phi^A(x^\mu), \partial_\nu \phi^A(x^\mu)). \tag{1.47}$$

As in the case of symplectic geometry, it is possible to obtain the conservation of multisymplecticity by either pulling back an $(n + 1)$-form from the Hamiltonian side, or by staying completely on the Lagrangian side.

In the symplectic case the canonical 1-form was defined on the cotangent bundle, the dual of the tangent bundle. Correspondingly, in multisymplectic geometry, we have the dual jet bundle, $J^1(Y)^*$, of fibre dimension $n \times N + 1$. This is defined as a vector bundle over Y whose fibre at $y \in Y_x$ is the set of affine maps $J^1(Y)_y \mapsto \Lambda^{n+1}(X)_x$, where $\Lambda^{n+1}(X)$ is the bundle of $(n+1)$-forms on X. Thus, a smooth section of the dual jet bundle is an affine bundle map of $J^1(Y)^*$ to $\Lambda^{n+1}(X)$ covering π_{XY} and takes the form

$$v^A{}_\mu \mapsto (p + p_A{}^\mu v^A{}_\mu)\mathrm{d}^{n+1}x, \tag{1.48}$$

where $\mathrm{d}^{n+1}x = \mathrm{d}x^0 \wedge \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n$ and $(p, p_A{}^\mu)$ are fibre coordinates on $J^1(Y)^*$. Here, p is referred to as the covariant Hamiltonian.

Then, we can define the canonical $(n+1)$-form on $J^1(Y)^*$ by

$$\Theta = p_A{}^\mu \mathrm{d}y^A \wedge \mathrm{d}^n x_\mu + p\mathrm{d}^{n+1}x, \tag{1.49}$$

and the canonical $(n+2)$-form on $J^1(Y)^*$ by

$$\Omega = -\mathrm{d}\Theta = \mathrm{d}y^A \wedge \mathrm{d}p_A{}^\mu \wedge \mathrm{d}^n x_\mu - \mathrm{d}p \wedge \mathrm{d}^{n+1}x, \tag{1.50}$$

which is both closed ($\mathrm{d}\Omega = 0$) and exact ($\Omega = \mathrm{d}\alpha$ for some $n+1$-form α).

Now, on the jet bundle $J^1(Y)$ we can introduce the Lagrangian $L(x^\mu, y^A, v^A{}_\mu)$ with the Lagrangian density $\mathcal{L} : J^1(Y) \mapsto \Lambda^{n+1}(X)$, which is an $(n+1)$-form given by

$$\mathcal{L}(\gamma) = L(x^\mu, y^A, v^A{}_\mu)\mathrm{d}^{n+1}x. \tag{1.51}$$

This Lagrangian density intrinsically defines a covariant Legendre transformation by vertical differentiation, i.e., the fibre preserving map over Y given by $J^1(Y) \mapsto J^1(Y)^*$. Let $\gamma, \gamma' \in J^1(Y)_y$, then

$$\mathbb{F}\mathcal{L}(\gamma) \cdot \gamma' = \mathcal{L}(\gamma) + \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} \mathcal{L}(\gamma + \epsilon(\gamma' - \gamma)). \tag{1.52}$$

The result of the Legendre transformation is that the fibre coordinates of $J^1(Y)^*$ are given by

$$p_A{}^\mu = \frac{\partial L}{\partial v^A{}_\mu} \text{ and } p = L - \frac{\partial L}{\partial v^A{}_\mu} v^A{}_\mu. \tag{1.53}$$

A multisymplectic manifold can then be defined, in some sense [30], as the pairing of the space Z (which is an appropriate dual of the jet bundle whose fibre over $y \in Y$ is given by

$$Z_y = \{z \in \Lambda_y | \imath_v \imath_w z = 0, \qquad \text{for all } v, w \in V_y Y\}, \tag{1.54}$$

where Λ_y is the fibre over $y \in Y$ of $\Lambda^{n+1}(Y)$ and $V_y Y = \{v \in T_y Y | T\pi_{XY}.v = 0\}$. Z can be roughly interpreted as the covariant cotangent bundle of the bundle Y) with its canonical $(n+2)$-form, i.e., (Z, Ω).

In a similar manner to the symplectic case, the Legendre transform can be used to define an $(n+1)$-form on $J^1(Y)$, known as the Cartan form, given by the pull-back of Θ, i.e.,

$$\Theta_{\mathcal{L}} = (\mathbb{F}\mathcal{L})^* \Theta. \tag{1.55}$$

The Cartan $(n+2)$-form then results from exterior differentiation,

$$\Omega_{\mathcal{L}} = -\mathrm{d}\Theta_{\mathcal{L}} = (\mathbb{F}\mathcal{L})^* \Omega. \tag{1.56}$$

Expressed in coordinates, these forms are given by

$$\Theta_{\mathcal{L}} = \frac{\partial L}{\partial v^A{}_\mu} \mathrm{d}y^A \wedge \mathrm{d}^n x_\mu + \left(L - \frac{\partial L}{\partial v^A{}_\mu} v^A{}_\mu \right) d^{n+1}x \tag{1.57}$$

and

$$\Omega_{\mathcal{L}} = \mathrm{d}y^A \wedge \mathrm{d}\left(\frac{\partial L}{\partial v^A{}_\mu} \right) \wedge \mathrm{d}^n x_\mu - \mathrm{d}\left[L - \frac{\partial L}{\partial v^A{}_\mu} v^A{}_\mu \right] \wedge \mathrm{d}^{n+1}x. \tag{1.58}$$

The alternate method of obtaining the Cartan $(n+1)$- and $(n+2)$-forms is to avoid the dual jet bundle altogether and stay completely on the Lagrangian side. I describe the key steps in this process below; for a more detailed derivation the reader is referred to [52].

Let U be a smooth manifold with (piecewise) smooth closed boundary. Then, define the set of smooth maps

$$\mathcal{C}^\infty = \{\phi : U \mapsto Y | \pi_{XY} \circ \phi : U \mapsto X \text{ is an embedding}\}, \tag{1.59}$$

where for each $\phi \in \mathcal{C}^\infty$, we set $\phi_X := \pi_{XY} \circ \phi$ and $U_X := \pi_{XY} \circ \phi(U)$ so that $\phi_X : U \mapsto U_X$ is a diffeomorphism, and let $\mathcal{C}$ be the closure of $\mathcal{C}^\infty$. The authors of [52] remark that U_X could be extended to the entire manifold X if their proofs were modified to handle the boundary ∂X.

Let $\mathcal{G}$ be a Lie group of π_{XY}-bundle automorphisms η_Y, covering diffeomorphisms η_X, with a Lie algebra $\mathfrak{g}$. Let an action $\Phi : \mathcal{G} \times \mathcal{C} \mapsto \mathcal{C}$ be defined by $\Phi(\eta_Y, \phi) = \eta_Y \circ \phi$. Then, the tangent space to the manifold $\mathcal{C}$ at a point ϕ in the set $T_\phi \mathcal{C}$ is defined by

$$\{V \in \mathcal{C}^\infty(X, TY) | \pi_{Y,TY} \circ V = \phi, T\pi_{XY} \circ V = V_X \text{ is a vector field on } X\}. \tag{1.60}$$

The vector field V on $T_\phi \mathcal{C}$ can be extended to the vector field $\mathcal{V}$ on $\mathcal{C}$ by fixing $v \in TY$ such that $V = v \circ (\phi \circ \phi_X^{-1})$ and letting $\mathcal{V}_\rho = v \circ (\rho \circ \rho_X^{-1})$. This gives the flow of $\mathcal{V}$ on $\mathcal{C}$ as $\Phi(\eta_Y^\lambda, \rho)$, where η_Y^λ, covering η_X^λ, is the flow of v.

Now, let the infinitesimal generators V and V_X be given by

$$V = \left.\frac{\mathrm{d}}{\mathrm{d}\lambda}\right|_{\lambda=0} \Phi(\eta_Y^\lambda, \phi) \text{ and } V_X = \left.\frac{\mathrm{d}}{\mathrm{d}\lambda}\right|_{\lambda=0} \eta_X^\lambda \circ \phi. \tag{1.61}$$

Then the action functional $\mathcal{S}$ on $\mathcal{C}$ is defined as

$$\mathcal{S}(\phi) = \int_{U_X} \mathcal{L}(j^1(\phi \circ \phi_X^{-1})), \qquad \text{for all } \phi \in \mathcal{C}, \tag{1.62}$$

and has an extremum when

$$\mathrm{d}\mathcal{S}_\phi \cdot V = \left.\frac{\mathrm{d}}{\mathrm{d}\lambda}\right|_{\lambda=0} \mathcal{S}(\Phi(\eta_Y^\lambda, \phi)) = 0. \tag{1.63}$$

The next step is to use V and V_X to write Eq. (1.63) as

$$\mathrm{d}\mathcal{S}_\phi \cdot V = \int_{U_X} \left.\frac{\mathrm{d}}{\mathrm{d}\lambda}\right|_{\lambda=0} \mathcal{L}(j^1(\Phi(\eta_Y^\lambda, \phi))) + \int_{\partial U_X} \imath_{V_X} \mathcal{L}(j^1(\phi \circ \phi_X^{-1})), \tag{1.64}$$

which results from splitting V into $V^v + V^h$, where $V^h = T(\phi \circ \phi_X^{-1}) \cdot V_X$ and $V^v = V - V^h$, and one application of the Cartan formula.

The Cartan $(n+1)$-form is then defined in Theorem 4.4 of [52], which I quote here.

Theorem 4.4. *Given a smooth Lagrangian density $\mathcal{L} : J^1(Y) \mapsto \Lambda^{n+1}(X)$, there exist a unique smooth section $D_{EL}\mathcal{L} \in C^\infty(Y'', \Lambda^{n+1}(X) \otimes T^*Y)$ and a unique differential form $\Theta_\mathcal{L} \in \Lambda(J^1(Y))$ such that for any $V \in T_\phi C$, and any open subset U_X such that $\bar{U}_X \cap \partial X = 0$,*

$$\mathrm{d}S_\phi \cdot V = \int_{U_X} D_{EL}(j^2(\phi \circ \phi_X^{-1})) \cdot V + \int_{\partial U_X} j^1(\phi \circ \phi_X^{-1}) * [\imath_{j^1(V)} \Theta_\mathcal{L}].$$

Furthermore,

$$D_{EL}\mathcal{L}(j^2(\phi \circ \phi_X^{-1})) \cdot V = j^1(\phi \circ \phi_X^{-1})^* [\imath_{j^1(V)} \Omega_\mathcal{L}] \text{ in } U_X.$$

In coordinates, the action of the Euler–Lagrange derivative D_{EL} on Y'' is given by

$$\begin{aligned} D_{EL}\mathcal{L}(j^2(\phi \circ \phi_X^{-1})) = \Bigg[&\frac{\partial L}{\partial y^A}(j^1(\phi \circ \phi_X^{-1})) - \frac{\partial^2 L}{\partial x^\mu \partial v^A{}_\mu}(j^1(\phi \circ \phi_X^{-1})) \\ &- \frac{\partial^2 L}{\partial y^B \partial v^A{}_\mu}(j^1(\phi \circ \phi_X^{-1})) \cdot (\phi \circ \phi_X^{-1})^B_{,\mu} \\ &- \frac{\partial^2 L}{\partial v^B{}_\nu \partial v^A{}_\mu}(j^1(\phi \circ \phi_X^{-1})) \cdot (\phi \circ \phi_X^{-1})^B_{,\mu\nu} \Bigg] \mathrm{d}y^A \wedge \mathrm{d}^{n+1}x, \end{aligned} \tag{1.65}$$

while the form $\Theta_{\mathcal{L}}$ matches the definition of the Cartan given in (4.9) [in [52]] *and has the coordinate expression*

$$\Theta_{\mathcal{L}} = \frac{\partial L}{\partial v^A{}_\mu} \mathrm{d}y^A \wedge \mathrm{d}^n x_\mu + \left(L - \frac{\partial L}{\partial v^A{}_\mu} v^A{}_\mu \right) \mathrm{d}^{n+1}x.$$

The Cartan $(n+2)$-form is, once again, given by minus the exterior derivative of the Cartan $(n+1)$-form, i.e., $\Omega_{\mathcal{L}} = -\mathrm{d}\Theta_{\mathcal{L}}$.

1.3.3 $\mathcal{E}$-$\mathcal{L}$ equations and the multisymplectic form formula

Now that we have the Cartan forms, we can use them to obtain the Euler–Lagrange equations and show the conservation of multisymplecticity. Using the definition of V above, if V is π_{XY}-vertical, then the first jet extension of V is the vector field on J^1Y given by

$$j^1V = \left(0, V^A, \frac{\partial V^A}{\partial x^\mu} + \frac{\partial V^A}{\partial y^B} v^B{}_\mu \right). \tag{1.66}$$

This allows Eq. (1.63) to be re-written as

$$\mathrm{d}\mathcal{S}_\phi \cdot V = \int_{U_X} -j^1(\phi \circ \phi_X^{-1})^* [\imath_{j^1V} \Omega_{\mathcal{L}}] + \int_{\partial U_X} j^1(\phi \circ \phi_X^{-1})^* [\imath_{j^1V} \Theta_{\mathcal{L}}], \tag{1.67}$$

and using a standard argument of the calculus of variations, the extremums of $\mathcal{S}$ occur when $j^1(\phi \circ \phi_X^{-1})^* [\imath_{j^1V} \Omega_{\mathcal{L}}] = 0$. This can be shown to occur when j^1V is replaced by $W \in TJ^1(Y)$, and so the set of solutions of the Euler–Lagrange equations can be expressed as

$$\mathcal{P} = \left\{ \phi \in \mathcal{C} \mid j^1(\phi \circ \phi_X^{-1})^* [\imath_W \Omega_{\mathcal{L}}] = 0 \quad \forall\, W \in TJ^1(Y) \right\}, \tag{1.68}$$

which in coordinates reads as: $(\phi \circ \phi_X^{-1})$ is a solution of the Euler–Lagrange equations if

$$\frac{\partial L}{\partial y^A}(j^1(\phi \circ \phi_X^{-1})) - \frac{\partial}{\partial x^\mu} \left(\frac{\partial L}{\partial v^A{}_\mu}(j^1(\phi \circ \phi_X^{-1})) \right) = 0 \text{ in } U_X. \tag{1.69}$$

Associated with $\mathcal{P}$ is the set of solutions of the first variation of the Euler–Lagrange equations, this consists of

$$\mathcal{F} = \left\{ V \in T_\phi \mathcal{C} \mid j^1(\phi \circ \phi_X^{-1})^* \mathfrak{L}_{j^1(V)} [\imath_W \Omega_{\mathcal{L}}] = 0, \quad \forall\, W \in TJ^1(Y) \right\}. \tag{1.70}$$

Now, let $\phi \in \mathcal{P}$, then for all V and W in $\mathcal{F}$ we have the *multisymplectic form formula*

$$\int_{\partial U_X} j^1(\phi \circ \phi_x^{-1})^* \left[\imath_{j^1(V)} \imath_{j^1(W)} \Omega_{\mathcal{L}} \right] = 0, \tag{1.71}$$

which can be written in coordinates as

$$\int_{\partial U_X} \left[\frac{\partial}{\partial y^B} \left(\frac{\partial L}{\partial v^A{}_\mu} \circ j^1(\phi) \right) (W^A V^B - W^B V^A) \right.$$
$$\left. + \frac{\partial}{\partial v^A{}_\nu} \left(\frac{\partial L}{\partial v^A{}_\mu} \circ j^1(\phi) \right) (W^A V^B_\nu - W^B_\nu V^A) \right] \mathrm{d}^n x_\mu = 0. \quad (1.72)$$

The authors of [52] claim that this is a generalisation of the multisymplectic conservation law (1.19) since it does not rely on the multi-Hamiltonian structure of the PDE.

1.3.4 Particle mechanics example

As an example of the Lagrangian approach, I give here the Cartan forms and Euler–Lagrange equations for particle mechanics.

Let $X = \mathbb{R}^2$ describe a single spatial dimension plus time, (x, t), and let $Y = \mathbb{R}^2 \times Q$, where Q has dimension n and describes the configuration of the system. The section $\phi : X \mapsto Y$ is then given by

$$\phi : (x, t) \mapsto (x, t, q^i(x, t)). \quad (1.73)$$

Also, the first jet bundle over Y is given by $J^1(Y) = \mathbb{R}^2 \times TQ$ with coordinates $(x, t, q^i, q^i_x, q^i_t)$, and the first jet of ϕ is

$$j^1(\phi) : (x, t) \mapsto (x, t, q^i(x, t), q^i_x(x, t), q^i_t(x, t)). \quad (1.74)$$

Thus the Cartan $(n+1)$- and $(n+2)$-forms may be written as

$$\Theta_{\mathcal{L}} = \frac{\partial L}{\partial q^i_x} \mathrm{d}q^i \wedge \mathrm{d}t + \frac{\partial L}{\partial q^i_t} \mathrm{d}q^i \wedge \mathrm{d}x + \left(L - \frac{\partial L}{\partial q^i_x} q^i_x - \frac{\partial L}{\partial q^i_t} q^i_t \right) \mathrm{d}x \wedge \mathrm{d}t \quad (1.75)$$

and

$$\Omega_{\mathcal{L}} = \mathrm{d}q^i \wedge \mathrm{d}\left(\frac{\partial L}{\partial q^i_x} \right) \wedge \mathrm{d}t + \mathrm{d}q^i \wedge \mathrm{d}\left(\frac{\partial L}{\partial q^i_t} \right) \wedge \mathrm{d}x - \mathrm{d}\left(L - \frac{\partial L}{\partial q^i_x} q^i_x - \frac{\partial L}{\partial q^i_t} q^i_t \right) \wedge \mathrm{d}x \wedge \mathrm{d}t. \quad (1.76)$$

Plugging the first jet of ϕ into Eq. (1.69), the Euler–Lagrange equations for a PDE with 1 time and 1 spatial dimension are obtained. They take the form

$$\frac{\partial}{\partial x}\left(\frac{\partial L}{\partial q^i_x} \right) + \frac{\partial}{\partial t}\left(\frac{\partial L}{\partial q^i_t} \right) - \frac{\partial L}{\partial q^i} = 0, \quad (1.77)$$

where $L = L(x, t, q^i, q^i_x, q^i_t)$, and are related to the multi-Hamiltonian PDE (1.6) though a Legendre transform.

1.4 Existing methods

In order to implement a symplectic or multisymplectic integrator, it is necessary to discretise the equations of motion while preserving their symplectic or multisymplectic structure. That is, a discrete analogue of the continuous symplectic or multisymplectic conservation law must be satisfied by the discretised equations of motion. There are currently several possible discretisation schemes which allow this, a selection of the most commonly used symplectic and multisymplectic discretisation schemes are listed below.

A common feature of all these discretisation schemes is that a grid is laid down on the independent variables (the independent variables are typically t for Hamiltonian ODEs and (t, x) for multi-Hamiltonian PDEs), where the grid points (or *nodes*) are usually taken to have equal spacing, although this is generally not required. A cell in this grid is the region defined by consecutive nodes, e.g., for a typical Hamiltonian ODE, cell i is the region $[i\Delta t, (i+1)\Delta t)$.

1.4.1 Finite differences

Methods based on finite differences use weighted sums over nearby nodes to approximate derivatives and to approximate the value of the solution at non-node locations for function evaluations. These weighted sums are usually written in a shorthand notation as a stencil. Given sufficient initial data, these stencils allow one to calculate the value of the solution on the nodes at the next level in time.

The weights used in these stencils are derived by considering Taylor expansions of a test function at nearby grid points. For example, on an equispaced grid, the Taylor expansions of the function $f(x)$ at neighbouring nodes are given by

$$\begin{aligned} f(x-\Delta x) &= f(x) - \Delta x f'(x) + \frac{1}{2!}(\Delta x)^2 f''(x) + O(\Delta x^3), \\ f(x+\Delta x) &= f(x) + \Delta x f'(x) + \frac{1}{2!}(\Delta x)^2 f''(x) + O(\Delta x^3). \end{aligned} \tag{1.78}$$

Taking the difference of these two Taylor expansions gives a second-order centred finite difference for the first derivative of $f(x)$,

$$f'(x) = \frac{1}{2\Delta x}(f(x+\Delta x) - f(x-\Delta x)) + O(\Delta x^2). \tag{1.79}$$

This is the most common finite difference method used in the literature, however, other common variants include one-sided stencils, wider stencils for increased accuracy and stencils on non-equispaced grids.

When applied in both time and space, finite difference methods typically lead to box schemes such as [2]

$$\boldsymbol{K} D_t M_x \boldsymbol{z} + \boldsymbol{L} D_x M_t \boldsymbol{z} = \nabla S(M_t M_x \boldsymbol{z}), \tag{1.80}$$

where D and M are stencils defining finite differences and averages, respectively. These methods are simple and effective over broad ranges of accuracy requirements and relatively simple geometries.

1.4.2 Runge–Kutta

Consider an autonomous ODE, $\partial_t \boldsymbol{z} = f(\boldsymbol{z}(t))$, where $\boldsymbol{z}(t) \in \mathbb{R}^n$. Then, on an equispaced grid (with spacing Δt), an r-stage Runge–Kutta (RK) discretisation of this ODE is a system of equations coupling the node values $\boldsymbol{z}_i$ and $\boldsymbol{z}_{i+1}$ to the stage values $\boldsymbol{Z}_{i,j}$ at r internal stages given by

$$\begin{aligned}
\boldsymbol{Z}_{i,j} &= \boldsymbol{z}_i + \Delta t \sum_{k=1}^{r} a_{jk} \partial_t \boldsymbol{Z}_{i,k}, \qquad \text{for } j = 1, \ldots, r, \\
\boldsymbol{z}_{i+1} &= \boldsymbol{z}_i + \Delta t \sum_{j=1}^{r} b_j \partial_t \boldsymbol{Z}_{i,j},
\end{aligned} \tag{1.81}$$

where the new variables $\partial_t \boldsymbol{Z}_{i,j}$ satisfy the ODE, i.e.,

$$\partial_t \boldsymbol{Z}_{i,j} = f(\boldsymbol{Z}_{i,j}), \tag{1.82}$$

and the coefficients b_j and a_{jk} are chosen to satisfy certain order conditions [16].

The coefficients b_j and a_{jk} can be written succinctly in what is referred to as a *Butcher tableau*:

$$\begin{array}{c|ccc}
c_1 & a_{11} & \cdots & a_{1r} \\
\vdots & \vdots & \ddots & \vdots \\
c_r & a_{r1} & \cdots & a_{rr} \\
\hline
 & b_1 & \cdots & b_r
\end{array} \tag{1.83}$$

where $c_i = \sum_j a_{ij}$ [16]. It has been shown [48, 67, 72] that these methods will preserve a symplectic conservation law when

$$b_i a_{ij} + b_j a_{ji} - b_i b_j = 0, \qquad \text{for all } i, j = 1, \ldots, r, \tag{1.84}$$

thus these methods are often referred to as symplectic RK methods.

The simplest example of a symplectic RK method is the *implicit midpoint* method, for which the Butcher tableau is given by

$$\begin{array}{c|c}
\frac{1}{2} & \frac{1}{2} \\
\hline
 & 1
\end{array}. \tag{1.85}$$

If the ODE, $\partial_t \boldsymbol{z} = f(\boldsymbol{z})$, is a Hamiltonian ODE and is discretised with the implicit midpoint method, then the stage values can be completely eliminated and the symplectic integrator can be written as

$$\boldsymbol{z}_{i+1} = \boldsymbol{z}_i + \Delta t f\left(\frac{\boldsymbol{z}_i + \boldsymbol{z}_{i+1}}{2}\right). \tag{1.86}$$

When a multi-Hamiltonian PDE (1.6) is discretised in time with one RK method and in space with another RK method, the result is a system of equations that has been shown to satisfy a discrete multisymplectic conservation law [63]. However, this system of equations does not, in general, form a well-defined integrator. This will be shown in Chapter 2, where Runge–Kutta discretisations of multi-Hamiltonian PDEs will be explored in more detail.

An example of when applying RK methods in time and space *does* result in a well-defined multisymplectic integrator is given in [2], where the authors apply the implicit midpoint method in time and space to the multi-Hamiltonian form of the KdV equation. The resulting method is equivalent to the well-known centred Preissman box scheme, which was introduced by Preissman in 1960 and is widely used in fluid mechanics. This method has been directly shown to satisfy a discrete multisymplectic conservation law [13].

1.4.3 Partitioned Runge–Kutta

Partitioned Runge–Kutta (PRK) methods are typically defined in the literature as a pair of RK discretisations with an equal number of internal stages, where the variables are partitioned into two partitions and each RK discretisation is applied to one of the partitions. However, this is only a subset of the class of PRK methods; I give a more general definition of a PRK method in Chapter 3. Nevertheless, I restrict the discussion here to the commonly used 2-partition PRK methods.

Let the variables $\boldsymbol{z}$ be partitioned into two partitions, $\boldsymbol{q}$ and $\boldsymbol{p}$. Then a 2-partition PRK discretisation of the non-autonomous system of ODEs,

$$\begin{aligned} \partial_t \boldsymbol{q} &= f(t, \boldsymbol{q}, \boldsymbol{p}), \\ \partial_t \boldsymbol{p} &= g(t, \boldsymbol{q}, \boldsymbol{p}), \end{aligned} \tag{1.87}$$

is given by

$$\begin{aligned} \boldsymbol{Q}_{i,j} &= \boldsymbol{q}_i + \Delta t \sum_{k=1}^{r} a_{jk} f(t_i + c_k \Delta t, \boldsymbol{Q}_{i,k}, \boldsymbol{P}_{i,k}), \\ \boldsymbol{q}_{i+1} &= \boldsymbol{q}_i + \Delta t \sum_{j=1}^{r} b_j f(t_i + c_j \Delta t, \boldsymbol{Q}_{i,j}, \boldsymbol{P}_{i,j}), \\ \boldsymbol{P}_{i,j} &= \boldsymbol{p}_i + \Delta t \sum_{k=1}^{r} \hat{a}_{jk} g(t_i + \hat{c}_k \Delta t, \boldsymbol{Q}_{i,k}, \boldsymbol{P}_{i,k}), \\ \boldsymbol{p}_{i+1} &= \boldsymbol{p}_i + \Delta t \sum_{j=1}^{r} \hat{b}_j g(t_i + \hat{c}_j \Delta t, \boldsymbol{Q}_{i,j}, \boldsymbol{P}_{i,j}), \end{aligned} \tag{1.88}$$

where a_{jk}, b_j, c_j, $\hat{a}_{jk}$, $\hat{b}_j$ and $\hat{c}_j$ can be written as a pair of Butcher tableaux:

$$\begin{array}{c|ccc} c_1 & a_{11} & \cdots & a_{1r} \\ \vdots & \vdots & \ddots & \vdots \\ c_r & a_{r1} & \cdots & a_{rr} \\ \hline & b_1 & \cdots & b_r \end{array} \qquad \begin{array}{c|ccc} \hat{c}_1 & \hat{a}_{11} & \cdots & \hat{a}_{1r} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{c}_r & \hat{a}_{r1} & \cdots & \hat{a}_{rr} \\ \hline & \hat{b}_1 & \cdots & \hat{b}_r \end{array} \tag{1.89}$$

where $c_i = \sum_j a_{ij}$ and $\hat{c}_i = \sum_j \hat{a}_{ij}$.

The benefits of using a PRK discretisation are not immediately apparent from Eq. (1.88). Indeed, Eq. (1.88) appears at first to be more complicated than Eq. (1.81). However, if the function f is independent of $\boldsymbol{q}$ and the function g is independent of $\boldsymbol{p}$, as is the case of separable Hamiltonian ODEs, then Eq. (1.88) becomes explicit.

For an autonomous Hamiltonian ODE discretised by a PRK method, the condition that the resulting system of equations satisfies a discrete symplectic conservation law is given by [1]

$$b_i\hat{a}_{ij} + \hat{b}_j a_{ji} - b_i\hat{b}_j = 0, \qquad \text{for } i, j = 1, \ldots, r. \tag{1.90}$$

Such PRK methods are often referred to as *symplectic PRK methods*.

The simplest example of a symplectic PRK method is the well-known *Störmer–Verlet* integrator, which has coefficients given by the following pair of tableaux

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & \frac{1}{2} & \frac{1}{2} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array} \qquad \begin{array}{c|cc} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}. \tag{1.91}$$

A special feature of the Störmer–Verlet integrator is that the internal stage variables, $\boldsymbol{P}_{i,1}$ and $\boldsymbol{P}_{i,2}$, are equal and can be relabelled as $\boldsymbol{p}_{i+\frac{1}{2}}$. Then, for an autonomous, separable Hamiltonian, the surrounding node variables, $\boldsymbol{p}_i$, can be eliminated in favour of the new variables on a staggered grid. This gives the following integrator:

$$\begin{aligned} \boldsymbol{p}_{i+1/2} &= \boldsymbol{p}_{i-1/2} + \Delta t g(\boldsymbol{q}_i), \\ \boldsymbol{q}_{i+1} &= \boldsymbol{q}_i + \Delta t f(\boldsymbol{p}_{i+1/2}), \end{aligned} \tag{1.92}$$

which is often called the *leapfrog* integrator for obvious reasons.

As with RK methods, PRK methods can be used to discretise a multi-Hamiltonian PDE in both time and space. It has been shown that discretisation of a multi-Hamiltonian PDE in time and space by PRK methods (with coefficients satisfying Eq. (1.90) and the same partitioning of the variables in time and space) gives a system of equations that formally satisfy a discrete multisymplectic conservation law [38]. However, such discretisations of multi-Hamiltonian PDEs suffer many of the same problems, with regards to forming a well-defined integrator, as RK discretisations.

In Chapter 3, I give a more general definition of a PRK discretisation based on an arbitrary partitioning of the variables. Furthermore, I give a simple set of conditions on the coefficients of these methods such that if they are applied independently in each dimension to a multi-Hamiltonian PDE, then the resulting system of equations satisfies a discrete multisymplectic conservation law.

1.4.4 Splitting methods

Often the vector field derived from a Hamiltonian contains terms that, while relatively simple to integrate individually, present difficulties when integrated together. For systems like this it is often possible to *split* the vector field into several pieces, where each piece is chosen to be simpler than the original vector field. These new vector fields may be simpler in one of the following two ways:

- the structure of the vector fields may be simpler, e.g., they may contain extra symmetries that are not in the original system;
- the vector fields may be easier to treat numerically, e.g., a fast Fourier transform may be applied.

Each part of the split system is then integrated separately and the parts are combined together in such a way that they preserve the structure of the flow of the original system [57]. This process is described more precisely for a symplectic splitting in the following example.

Consider the abstract Hamiltonian system $\boldsymbol{z}_t = \boldsymbol{J}^{-1}\nabla_{\boldsymbol{z}} H(\boldsymbol{z})$, with a time-$\Delta t$ flow given by

$$\boldsymbol{z}(\Delta t) = \exp(\Delta t \boldsymbol{J}^{-1}\nabla_{\boldsymbol{z}} H(\boldsymbol{z}))(\boldsymbol{z}(0)). \tag{1.93}$$

Suppose, that this vector field can be split into several simpler vector fields,

$$\boldsymbol{z}_t = \boldsymbol{J}^{-1}\nabla_{\boldsymbol{z}}(H^{(1)}(\boldsymbol{z}) + \cdots + H^{(N)}(\boldsymbol{z})), \tag{1.94}$$

where each of the subsystems $\boldsymbol{z}_t = \boldsymbol{J}^{-1}\nabla_{\boldsymbol{z}} H^{(j)}(\boldsymbol{z})$, $j = 1, \ldots, N$ can be solved exactly. Let the time-Δt solution operator of subsystem j be denoted by $\Psi^{(j)}(\Delta t)$, i.e., $\boldsymbol{z}(\Delta t) = \Psi^{(j)}(\Delta t)\boldsymbol{z}(0)$ satisfies the flow of subsystem j. Then, since the flow of each subsystem is a symplectic map and the composition of symplectic maps is again symplectic, the composite map

$$\bar{\Psi}(\Delta t) := \Psi^{(N)}(\Delta t) \circ \cdots \circ \Psi^{(1)}(\Delta t) \tag{1.95}$$

is symplectic. Furthermore, by the Baker–Campbell–Hausdorff (BCH) Theorem, it is a first order approximation of the exact flow map. Therefore, $\bar{\Psi}(\Delta t)$ is a first order symplectic integrator.

By varying the lengths of the solution intervals of the subsystems, and introducing more elaborate compositions, one can obtain symmetric, symplectic maps of arbitrary accuracy [79]. Moreover, the exact flow of each subsystem does not have to be available, any consistent symplectic integrator will do. Symplectic splitting methods are the most effective way of obtaining fast, efficient, explicit symplectic methods. For Poisson systems, they are often the only hope of constructing Poisson integrators. For a thorough account of splitting methods, see the review article by McLachlan and Quispel [57].

Until very recently, splitting methods had not been considered for multi-Hamiltonian PDEs as splitting methods are global constructions (since the PDE along with its boundary conditions are treated as an infinite dimensional ODE), whereas multisymplecticity is a local property and is independent of boundary conditions. Despite this, splitting methods for multi-Hamiltonian PDEs may be constructed which do satisfy a discrete multisymplectic conservation law [66]. This will be discussed in more detail in Section 4.4, with particular attention being given to the NLS equation.

Note that the question of multisymplecticity of splitting methods may be important for multisymplectic systems with nonlinear operators $\boldsymbol{K}(z)$ and $\boldsymbol{L}(z)$, since RK methods do not preserve such structures. This was briefly noted in [26].

1.4.5 Discrete variational integrators

Discrete variational integrators take a different approach from the methods listed above. Instead of working with a Hamiltonian or multi-Hamiltonian system of equations, a discrete analogue of the Lagrangian is constructed, along with discrete analogues of the apparatus of differential geometry. In particular, a discrete analogue of Hamilton's variational principle is constructed, which leads to a discretisation of the Euler–Lagrange equations. These discrete Euler–Lagrange equations are referred to as a discrete variational integrator.

The discrete Lagrangian is defined by $L : Q \times Q \mapsto \mathbb{R}$, where $(q_1, q_0) \in Q \times Q$ corresponds to $(q_1 - q_0)/\Delta t \in TQ$. Similarly, the discrete Euler–Lagrange (DEL) equations are given by

$$\frac{\partial L}{\partial q_1}(q_1, q_0) + \frac{\partial L}{\partial q_2}(q_2, q_1) = 0. \tag{1.96}$$

The discrete action corresponding to the discrete Lagrangian is defined by $S \equiv \sum_{k=1}^{n} L(q_k, q_{k-1})$. By the discrete analogue of Hamilton's variational principle, taking variations of the discrete action while holding the end points fixed allows one to calculate the discrete flow, $F : Q \times Q \mapsto Q \times Q$, such that $F(q_1, q_0) = (q_2, q_1)$. By looking at $\mathrm{d}S$ for variations δq that do not fix the end points, one can define two discrete 1-forms,

$$\theta_L^- = \frac{\partial L}{\partial q_0}(q_1, q_0)\delta q_0, \tag{1.97}$$

and

$$\theta_L^+ = \frac{\partial L}{\partial q_1}(q_1, q_0)\delta q_1. \tag{1.98}$$

The discrete 2-form is then given by

$$\omega_L = \mathrm{d}\theta_L^- = -\mathrm{d}\theta_L^+. \tag{1.99}$$

Thus, the discrete conservation of symplecticity can be written as

$$F^*\omega_L = \omega_L. \tag{1.100}$$

For multisymplectic PDEs, the Lagrangian approach is further complicated by the extension of the base manifold to higher than one dimension. In the simplest case, the base manifold is just 1 time and 1 spatial dimension, which is discretised as a rectangular grid of points, $X = (i, j)$, and the fibres of the configuration space, Y, are given by $Q = (q^0, \ldots, q^n)$. As in the continuous case, we also have a section, ϕ, mapping X (or a closed subset U) into Y. The situation now becomes complicated as we need to construct the discrete analogue of the jet bundle and we have a choice as to how we do this.

In [52], the authors label a triplet in X as a *triangle*, $\triangle = ((i, j), (i, j+1), (i+1, j+1))$, which defines a triplet in the fibres of Y given by $(y_{i,j}, y_{i,j+1}, y_{i+1,j+1})$. The first jet extension of ϕ is then defined to be the map $j^1\phi \mapsto J^1Y$ given by

$$j^1\phi(\triangle) \equiv (\triangle, \phi(\triangle^1), \phi(\triangle^2), \phi(\triangle^3)), \tag{1.101}$$

where $\triangle^i$ is component i of the triangle. The purpose of the triangle notation is that the discrete analogue of the derivatives, $\partial_\nu \phi^A(x^\mu)$, can be represented by a linear combination of the components of the triangle.

For example,

$$\frac{\partial \phi}{\partial t}(\bar{\triangle}) = \frac{1}{\Delta t}(\phi(\triangle^2) - \phi(\triangle^1)), \tag{1.102}$$

$$\frac{\partial \phi}{\partial x}(\bar{\triangle}) = \frac{1}{\Delta x}(\phi(\triangle^3) - \phi(\triangle^2)), \tag{1.103}$$

where $\bar{\triangle}$ is the centre of the triangle.

However, it should be pointed out that the choice to use triplets is somewhat arbitrary as other equally valid discretisations can be obtained from different linear combinations of the components of the triangle or by using other shapes, such as rectangles or hexagons.

Sticking with triangles, a discrete Lagrangian for each $\triangle$ can be constructed as $L_\triangle(y_{\triangle^1}, y_{\triangle^2}, y_{\triangle^3})$, giving the discrete action as

$$S(\phi) \equiv \sum_{\triangle;\triangle\subseteq U} L \circ j^1\phi(\triangle), \tag{1.104}$$

where U is regular. From the action, the discrete Euler–Lagrange field (DELF) equations can be obtained by differentiating the action with respect to each coordinate in Y. That is, the DELF equations are given by

$$\sum_{l;\triangle;(i,j)\in\triangle^l} \frac{\partial L_\triangle}{\partial y^l}(y_{\triangle^1}, y_{\triangle^2}, y_{\triangle^3}) = 0, \tag{1.105}$$

for all $(i,j) \in U \backslash \partial U$.

Now, taking variations of S, the boundary of U gives rise to the 1-form

$$\Theta^1_L(y_{\triangle^1}, y_{\triangle^2}, y_{\triangle^3}) \cdot (v_{\triangle^1}, v_{\triangle^2}, v_{\triangle^3}) \equiv \frac{\partial L}{\partial y^1}(y_{\triangle^1}, y_{\triangle^2}, y_{\triangle^3}) \cdot (v_{\triangle^l}, 0, 0) \tag{1.106}$$

with Θ^2_L and Θ^3_L defined similarly. The discrete Cartan form can then be written as

$$\theta_L(\phi) \cdot V \equiv \sum_{\triangle;\triangle\cap\partial U\neq\emptyset} \left(\sum_{l;\triangle^l\in\partial U} [(j^1\phi)^*(\imath_{j^1V}\Theta^l_L)](\triangle) \right). \tag{1.107}$$

Defining $\Omega^l_L = -\mathrm{d}\Theta^l_L$ and noting that $\mathrm{dd}S = 0$ immediately gives the *discrete multisymplectic form formula*,

$$\sum_{\triangle;\triangle\cap\partial U\neq\emptyset} \left(\sum_{l;\triangle^l\in\partial U} [(j^1\phi)^*(\imath_{j^1V}\imath_{j^1W}\Omega^l_L)](\triangle) \right) = 0, \tag{1.108}$$

which indicates that multisymplecticity is conserved by the integrator. For further details, the reader is referred to [52].

This approach has lead to a number of high-order multisymplectic integrators with properties such as discrete conservation of momentum, long-time accuracy of the solution and energy, and the preservation of statistical quantities [50].

1.5 Advantages and usage

As mentioned earlier, symplectic and multisymplectic integrators have enjoyed a large amount of attention in recent years. The reason for this attention is due to the many advantages that symplectic and multisymplectic integrators have over various other types of geometric integrators.

The defining feature of symplectic and multisymplectic integrators is that the integrator is a diffeomorphism from a symplectic or multisymplectic manifold to itself and thus the solutions of these integrators lie precisely on the same symplectic or multisymplectic manifold as the initial conditions. This has a number of important consequences, which I describe below.

A generalisation of Liouville's theorem states that for a Hamiltonian system with a divergence-free vector field, the flow is volume-preserving. This is a direct result of the conservation of symplecticity and as such, symplectic and multisymplectic integrators preserve volume in phase space.

It was shown recently that the preservation of a discrete multisymplectic conservation law implies the conservation of potential vorticity [39]. The conservation of potential vorticity is important in applications such as weather prediction.

For the symplectic integration of an integrable or nearly integrable Hamiltonian ODE, the difference between the numerical solution and the exact solution is exponentially small in the step size and grows linearly with time [34]. This result is achieved through a technique known as backward error analysis, whereby the sequence, $\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2, \ldots$ (which is produced by applying a numerical integrator with step size Δt to the Hamiltonian ODE, $\partial_t \boldsymbol{z} = f(\boldsymbol{z})$), coincides with the exact solution of a modified Hamiltonian ODE of the form

$$\partial_t \tilde{\boldsymbol{z}} = f_{\Delta t}(\tilde{\boldsymbol{z}}) = f(\tilde{\boldsymbol{z}}) + \Delta t f_2(\tilde{\boldsymbol{z}}) + (\Delta t)^2 f_3(\tilde{\boldsymbol{z}}) + \cdots , \tag{1.109}$$

i.e., $\boldsymbol{z}_n = \tilde{\boldsymbol{z}}(n\Delta t)$. The equivalent result for multi-Hamiltonian PDEs is yet to be proven in general, however significant steps in this direction have been taken [58, 59, 61].

A consequence of the numerical solution being the exact solution of some nearby Hamiltonian ODE or multi-Hamiltonian PDE is that the energy and momentum of the nearby system, which is close to the energy and momentum of the original system, is exactly conserved. This has led to the discovery of so-called *approximate conservation laws*. For Hamiltonian ODEs, the total energy in the system is approximately conserved by symplectic integrators, while for multi-Hamiltonian PDEs, both the energy and the momentum are approximately locally conserved by multisymplectic integrators. These approximate conservation laws have received much attention in the literature [8, 13, 58, 63].

Since the linear error growth in symplectic and multisymplectic integrators is exponentially small in the step size, a large time step (relative to time steps used in other numerical integrators) may be used while retaining a relatively small error. Furthermore, due to a lack of spurious non-physical effects, the numerical solution of symplectic and multisymplectic integrators have excellent long-time stability, which allows very lengthy simulations to be performed [34]. Moreover, symplectic and multisymplectic integrators often require fewer function evaluations than other integrators of the same order (in some cases the integrator may even be explicit) and thus require less processing time for the same amount of accuracy in the solution as is achievable with other integrators.

As well as the near preservation of invariant tori due to the exponentially small linear error growth mentioned above, it has been shown [34, 68] that a discrete version of Kolmogorov-Arnold-Moser (KAM) theory applies to the symplectic integration of integrable or near-integrable Hamiltonian ODEs with KAM tori.

Lastly, it should be recalled that a multisymplectic conservation law may be reduced to a symplectic conservation law by integration over space with suitable boundary conditions. Thus, any results for integrators that preserve a symplectic conservation law also apply in a global sense to integrators that preserve a discrete multisymplectic conservation law. However, multisymplectic integrators may possess local properties that are, as yet, unknown and may not correspond to any property of symplectic integrators.

Symplectic integrators have been applied to a vast variety of Hamiltonian ODEs with extraordinary results. These systems range in complexity from very low degree of freedom systems, such as the double pendulum or harmonic oscillator, to very high degree of freedom systems, such as molecular or stellar dynamics (e.g., the authors of [23] used a symplectic integrator to carry out a one million year simulation of the whole solar system including the moons and obliquity of Earth and Mars in order to study the climatic variations of Earth and Mars over the past one million years. Their results showed a very high correlation between ice ages and the Earth's orbit over this period).

Multisymplectic integrators have been applied to remarkably fewer multi-Hamiltonian PDEs. This is due, in part, to the field of multisymplectic integration being relatively new and to an incomplete classification of the PDEs which can be written as a well-defined first-order system in the form of a multi-Hamiltonian PDE (1.6).

Some of the equations that have had multisymplectic integrators established for them are the Benjamin-Bona-Mahony equation [65], the Boussinesq equation [17, 65], the Gross-Pitaevskii equation [41], the Kadomtsev-Petviashvili equation [51], the KdV equation [2, 65], the Klein-Gordon equation [8], the Landau-Lifshitz equation [26], Maxwell's equations [69], the nonlinear Schrödinger equation [60, 63, 18, 41, 65, 66], the nonlinear wave equation [60, 12, 59, 63, 65], the Padé–II equation [25, 65] the Sine-Gordon equation [60, 63], and the Zakharov-Kuznetsov equation [14].

Chapter 2

Runge–Kutta Discretisations

A Runge–Kutta (RK) method for the discretisation of an ODE was defined in section 1.4.2 as a system of equations coupling adjacent node values to a number of internal stage values. These equations are given by Eqs. (1.81) and (1.82), and contain a number of coefficients that can be written concisely in a Butcher tableau (Eq. (1.83)). These equations define a map, $\boldsymbol{z}_i \mapsto \boldsymbol{z}_{i+1}$, which has an accuracy dependent on the time step, the number of internal stages and the coefficients in the Butcher tableau. Furthermore, if the RK method is applied to a Hamiltonian ODE and the coefficients satisfy Eq. (1.84), then the map is symplectic, i.e., Eqs. (1.81) and (1.82) satisfy the discrete analogue of the symplectic conservation law for that ODE. Such RK methods are often referred to as symplectic RK methods.

Many of the higher order RK methods are based on the concept of collocation, whereby the coefficients in the RK method are determined by a collocation polynomial, which passes through the node $\boldsymbol{z}_i$ and has derivatives matching the ODE at a number of quadrature points (i.e., the c_i values) [33]. In order to achieve the highest possible order from a RK method based on collocation, it is necessary that the quadrature points be distinct. The three types of quadrature that give rise to the highest order RK methods are Gaussian, Radau and Lobatto. For an r-stage RK method, they are:

- Gaussian quadrature.
  Gaussian RK (GRK) methods have quadrature points that are the zeros of the shifted Legendre polynomial of degree r,
  $$P_r(x) = \frac{1}{r!}\frac{\mathrm{d}^r}{\mathrm{d}x^r}\left(x^r(x-1)^r\right), \tag{2.1}$$
  which are in the range $(0, 1)$. These methods have order $2r$.
- Radau quadrature.
  There are two types of RK methods based on Radau quadrature, Radau left and

Radau right, both of which have order $2r-1$. The quadrature points of the Radau left method are the zeros of

$$\frac{\mathrm{d}^{r-1}}{\mathrm{d}x^{r-1}}\left(x^r(x-1)^{r-1}\right), \tag{2.2}$$

and are in the range $[0,1)$, with $c_1=0$. Whereas, the quadrature points of the Radau right method are the zeros of

$$\frac{\mathrm{d}^{r-1}}{\mathrm{d}x^{r-1}}\left(x^{r-1}(x-1)^{r}\right), \tag{2.3}$$

and are in the range $(0,1]$, with $c_r=1$.

- Lobatto quadrature.
  The quadrature points of RK methods based on Lobatto quadrature are the zeros of
  $$\frac{\mathrm{d}^{r-2}}{\mathrm{d}x^{r-2}}\left(x^{r-1}(x-1)^{r-1}\right), \tag{2.4}$$
  and are in the range $[0,1]$, with $c_1=0$ and $c_r=1$. These methods (called Lobatto IIIA methods for historical reasons [34]) are of order $2r-2$, but have an advantage over the above methods; while the collocation polynomials of each of the above methods are continuous over a range of cells, only the Lobatto collocation polynomials are smooth over a range of cells (the others are piecewise smooth).

  A different type of collocation is known as discontinuous collocation, where the collocation polynomial does not, in general, pass through the node $\boldsymbol{z}_i$. It has been shown that discontinuous collocation with r stages is equivalent to a RK method with $r-2$ stages [34]. Using discontinuous collocation with the Lobatto quadrature points gives rise to the Lobatto IIIB class of RK methods, which, when combined with Lobatto IIIA methods, form an important class of PRK methods as shall be seen in Chapter 3.

For a RK method based on collocation, the coefficients a_{ij} and b_j in the method satisfy the relations [16]:

$$\begin{aligned} B(\xi): &\quad \sum_{i=1}^{r} b_i c_i^{k-1} = \frac{1}{k}, &&\text{for } k\le\xi,\\ C(\xi): &\quad \sum_{j=1}^{r} a_{ij} c_j^{k-1} = \frac{1}{k}c_i^k, &&\text{for } i=1,\dots,r \text{ and } k\le\xi, \end{aligned} \tag{2.5}$$

for $\xi=r$. Now, the Gaussian and Lobatto quadrature points are symmetric about $\frac{1}{2}$. This leads to the following lemmas, which are proven in Appendix A.1 and Appendix A.2.

Lemma 2.0.1. *The coefficients, b_i, of RK methods based on collocation at Gaussian or Lobatto quadrature points satisfy $b_i=b_{r+1-i}$.*

Lemma 2.0.2. *The matrix $\boldsymbol{A}$ of the coefficients a_{ij} in the Runge–Kutta method satisfies*

$$(\boldsymbol{A}^{-1}\mathbf{1})_i = (-1)^{r-1} \prod_{j \neq i} (c_i - c_j) / \prod_j c_j, \tag{2.6}$$

if the matrix $\boldsymbol{A}$ is invertible.

The focus of this chapter is primarily on the discretisation of multi-Hamiltonian PDEs by applying Gaussian Runge–Kutta methods in space and time such that a discrete multisymplectic conservation law is satisfied.

As mentioned in Section 1.2.1, the concatenation of a discretisation conserving ω in time with a discretisation conserving κ in space does not necessarily give a multisymplectic discretisation of the multi-Hamiltonian PDE (1.6) since the conservation of ω in time is global in space and the conservation of κ in space is global in time while the discrete conservation of multisymplecticity is local in both time and space. Fortunately, for RK discretisations with coefficients satisfying Eq. (1.84), this concatenation does result in a system of equations that satisfy a discrete multisymplectic conservation law [63], as is shown in the next section for the nonlinear wave equation. (A general proof for arbitrary multi-Hamiltonian PDEs and general PRK discretisations (including RK) with coefficients satisfying a symplecticity condition is given in Section 3.2.)

2.1 The multisymplectic conservation law

It has previously been shown [63] that applying RK discretisations in time and space to a multi-Hamiltonian PDE results in a system of equations that satisfy a discrete multisymplectic conservation law. The example used in [63] was the nonlinear wave equation and so, in this chapter, I shall also use the nonlinear wave equation as the primary example. The nonlinear wave equation can be written as a multi-Hamiltonian PDE with

$$\boldsymbol{z} = \begin{bmatrix} u \\ v \\ w \end{bmatrix}, \quad \boldsymbol{K} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{L} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \tag{2.7}$$

and $S(\boldsymbol{z}) = -V(u) + \frac{1}{2}(w^2 - v^2)$.

Consider the nonlinear wave equation discretised by an s-stage GRK discretisation in time and an r-stage GRK discretisation in space, where the notation of Section 1.4.2 is used for the GRK discretisations, with raised indices for the time discretisation and lowered indices for the spatial discretisation, as shown in Figure 2.1.

The fully discretised system of equations for the cell with a lower left corner at

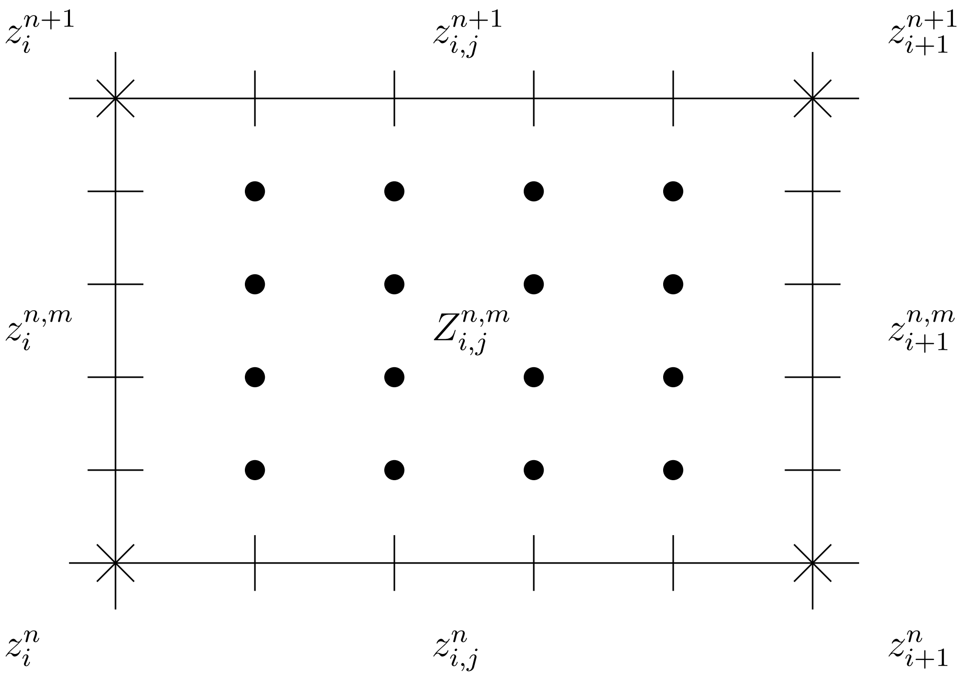


Figure 2.1: A single cell demonstrating the labelling convention for a Runge–Kutta discretisation in space and time.

$(n\Delta t, i\Delta x)$ is given by

$$\begin{aligned}
U_{i,j}^{n,m} &= u_i^{n,m} + \Delta x \sum_{k=1}^{r} a_{jk} \partial_x U_{i,k}^{n,m}, & u_{i+1}^{n,m} &= u_i^{n,m} + \Delta x \sum_{k=1}^{r} b_k \partial_x U_{i,k}^{n,m},\\
W_{i,j}^{n,m} &= w_i^{n,m} + \Delta x \sum_{k=1}^{r} a_{jk} \partial_x W_{i,k}^{n,m}, & w_{i+1}^{n,m} &= w_i^{n,m} + \Delta x \sum_{k=1}^{r} b_k \partial_x W_{i,k}^{n,m},\\
U_{i,j}^{n,m} &= u_{i,j}^{n} + \Delta t \sum_{l=1}^{s} \tilde{a}_{ml} \partial_t U_{i,j}^{n,l}, & u_{i,j}^{n+1} &= u_{i,j}^{n} + \Delta t \sum_{l=1}^{s} \tilde{b}_l \partial_t U_{i,j}^{n,l},\\
V_{i,j}^{n,m} &= v_{i,j}^{n} + \Delta t \sum_{l=1}^{s} \tilde{a}_{ml} \partial_t V_{i,j}^{n,l}, & v_{i,j}^{n+1} &= v_{i,j}^{n} + \Delta t \sum_{l=1}^{s} \tilde{b}_l \partial_t V_{i,j}^{n,l},
\end{aligned} \tag{2.8}$$

where $1 \le j \le r$, $1 \le m \le s$ and the terms $\partial_x U_{i,j}^{n,m}$, $\partial_x W_{i,j}^{n,m}$, $\partial_t U_{i,j}^{n,m}$ and $\partial_t V_{i,j}^{n,m}$ satisfy the PDE, i.e.,

$$\begin{aligned}
\partial_t U_{i,j}^{n,m} &= V_{i,j}^{n,m},\\
\partial_x U_{i,j}^{n,m} &= W_{i,j}^{n,m},\\
\partial_t V_{i,j}^{n,m} - \partial_x W_{i,j}^{n,m} &= -V'(U_{i,j}^{n,m}).
\end{aligned} \tag{2.9}$$

From $\omega = \frac{1}{2}\boldsymbol{K}\mathrm{d}z \wedge \mathrm{d}z$ and $\kappa = \frac{1}{2}\boldsymbol{L}\mathrm{d}z \wedge \mathrm{d}z$, it can be seen that the continuous multi-symplectic conservation law for the nonlinear wave equation is given by

$$\partial_t(\mathrm{d}u \wedge \mathrm{d}v) - \partial_x(\mathrm{d}u \wedge \mathrm{d}w) = 0, \tag{2.10}$$

which is satisfied since

$$\begin{aligned}\partial_t(\mathrm{d}u \wedge \mathrm{d}v) - \partial_x(\mathrm{d}u \wedge \mathrm{d}w) &= \mathrm{d}u \wedge \partial_t \mathrm{d}v - \mathrm{d}u \wedge \partial_x \mathrm{d}w \\ &= -V''(u)\mathrm{d}u \wedge \mathrm{d}u \\ &= 0.\end{aligned} \tag{2.11}$$

A discrete version of the multisymplectic conservation law is obtained by approximating the time derivative by $\partial_t(\mathrm{d}u \wedge \mathrm{d}v) = \frac{1}{\Delta t}(\mathrm{d}u_{i,j}^{n+1} \wedge \mathrm{d}v_{i,j}^{n+1} - \mathrm{d}u_{i,j}^{n} \wedge \mathrm{d}v_{i,j}^{n})$ and approximating the spatial derivative by $\partial_x(\mathrm{d}u \wedge \mathrm{d}w) = \frac{1}{\Delta x}(\mathrm{d}u_{i+1}^{n,m} \wedge \mathrm{d}w_{i+1}^{n,m} - \mathrm{d}u_{i}^{n,m} \wedge \mathrm{d}w_{i}^{n,m})$, where the terms $\mathrm{d}u_{i,j}^{n}$, $\mathrm{d}v_{i,j}^{n}$, $\mathrm{d}u_{i}^{n,m}$ and $\mathrm{d}w_{i}^{n,m}$ satisfy the first variation of Eq. (2.8). The result is [63]:

$$\Delta x \sum_{j=1}^{r} b_j(\mathrm{d}u_{i,j}^{n+1} \wedge \mathrm{d}v_{i,j}^{n+1} - \mathrm{d}u_{i,j}^{n} \wedge \mathrm{d}v_{i,j}^{n}) - \Delta t \sum_{m=1}^{s} \tilde{b}_m(\mathrm{d}u_{i+1}^{n,m} \wedge \mathrm{d}w_{i+1}^{n,m} - \mathrm{d}u_{i}^{n,m} \wedge \mathrm{d}w_{i}^{n,m}) = 0, \tag{2.12}$$

which is an approximation to the integral over the region $n\Delta t \leq t < (n+1)\Delta t$, $i\Delta x \leq x < (i+1)\Delta x$ (i.e., one cell) of the continuous multisymplectic conservation law.

Theorem 2.1.1. *The system of equations (Eqs. (2.8) and (2.9)) obtained by discretising the nonlinear wave equation in space by an r-stage symplectic GRK method and in time by an s-stage symplectic GRK method satisfies the discrete multisymplectic conservation law given by Eq. (2.12).*

Proof. Expanding the portion of Eq. (2.12) that corresponds to the term $\partial_t(\mathrm{d}u \wedge \mathrm{d}v)$ in the continuous multisymplectic conservation law gives

$$\begin{aligned}
&\mathrm{d}u_{i,j}^{n+1} \wedge \mathrm{d}v_{i,j}^{n+1} - \mathrm{d}u_{i,j}^{n} \wedge \mathrm{d}v_{i,j}^{n} \\
&\quad = \mathrm{d}u_{i,j}^{n} \wedge \Delta t \sum_{m=1}^{s} \tilde{b}_m \partial_t \mathrm{d}V_{i,j}^{n,m} + \Delta t \sum_{l=1}^{s} \tilde{b}_l \partial_t \mathrm{d}U_{i,j}^{n,l} \wedge \mathrm{d}v_{i,j}^{n} \\
&\qquad\qquad + (\Delta t)^2 \sum_{l=1}^{s}\sum_{m=1}^{s} \tilde{b}_l \tilde{b}_m \wedge \partial_t \mathrm{d}U_{i,j}^{n,m} \partial_t \mathrm{d}V_{i,j}^{n,m} \\
&\quad = \Delta t \sum_{m=1}^{s} \tilde{b}_m(\mathrm{d}U_{i,j}^{n,m} - \Delta t \sum_{l=1}^{s} \tilde{a}_{ml} \partial_t \mathrm{d}U_{i,j}^{n,l}) \wedge \partial_t \mathrm{d}V_{i,j}^{n,m} \\
&\qquad\qquad + \Delta t \sum_{l=1}^{s} \tilde{b}_l \partial_t \mathrm{d}U_{i,j}^{n,l} \wedge (\mathrm{d}V_{i,j}^{n,l} - \Delta t \sum_{m=1}^{s} \tilde{a}_{lm} \partial_t \mathrm{d}V_{i,j}^{n,m}) \\
&\qquad\qquad + (\Delta t)^2 \sum_{l=1}^{s}\sum_{m=1}^{s} \tilde{b}_l \tilde{b}_m \partial_t \mathrm{d}U_{i,j}^{n,l} \wedge \partial_t \mathrm{d}V_{i,j}^{n,m} \\
&\quad = \Delta t \sum_{m=1}^{s} \tilde{b}_m(\mathrm{d}U_{i,j}^{n,m} \wedge \partial_t \mathrm{d}V_{i,j}^{n,m} + \partial_t \mathrm{d}U_{i,j}^{n,m} \wedge \mathrm{d}V_{i,j}^{n,m}) \\
&\qquad\qquad + (\Delta t)^2 \sum_{l=1}^{s}\sum_{m=1}^{s} (-\tilde{b}_m \tilde{a}_{ml} - \tilde{b}_l \tilde{a}_{lm} + \tilde{b}_l \tilde{b}_m) \partial_t \mathrm{d}U_{i,j}^{n,l} \wedge \partial_t \mathrm{d}V_{i,j}^{n,m}.
\end{aligned} \tag{2.13}$$

When the coefficients of the GRK discretisation satisfy Eq. (1.84), the last term is zero. Also, $\partial_t \mathrm{d}U_{i,j}^{n,m} \wedge \mathrm{d}V_{i,j}^{n,m} = \mathrm{d}V_{i,j}^{n,m} \wedge \mathrm{d}V_{i,j}^{n,m} = 0$ by the first line of Eq. (2.9). This leaves

$$\mathrm{d}u_{i,j}^{n+1} \wedge \mathrm{d}v_{i,j}^{n+1} - \mathrm{d}u_{i,j}^{n} \wedge \mathrm{d}v_{i,j}^{n} = \Delta t \sum_{m=1}^{s} \tilde{b}_m (\mathrm{d}U_{i,j}^{n,m} \wedge \partial_t \mathrm{d}V_{i,j}^{n,m}). \tag{2.14}$$

Similarly,

$$\mathrm{d}u_{i+1}^{n,m} \wedge \mathrm{d}w_{i+1}^{n,m} - \mathrm{d}u_{i}^{n,m} \wedge \mathrm{d}w_{i}^{n,m} = \Delta x \sum_{j=1}^{r} b_j (\mathrm{d}U_{i,j}^{n,m} \wedge \partial_x \mathrm{d}W_{i,j}^{n,m}). \tag{2.15}$$

If one multiplies Eq. (2.15) by $\Delta t \tilde{b}_m$ and sums over m, multiplies Eq. (2.14) by $\Delta x b_j$ and sums over j, then subtracts the former from the latter, one finds that

$$\begin{aligned}
\Delta x \sum_{j=1}^{r} b_j (\mathrm{d}u_{i,j}^{n+1} \wedge \mathrm{d}v_{i,j}^{n+1} - \mathrm{d}u_{i,j}^{n} \wedge \mathrm{d}v_{i,j}^{n}) - \Delta t \sum_{m=1}^{s} \tilde{b}_m (\mathrm{d}u_{i+1}^{n,m} \wedge \mathrm{d}w_{i+1}^{n,m} - \mathrm{d}u_{i}^{n,m} \wedge \mathrm{d}w_{i}^{n,m})& \\
= \Delta t \Delta x \sum_{j=1}^{r} \sum_{m=1}^{s} \tilde{b}_m b_j (\mathrm{d}U_{i,j}^{n,m} \wedge \partial_t \mathrm{d}V_{i,j}^{n,m} - \mathrm{d}U_{i,j}^{n,m} \wedge \partial_x \mathrm{d}W_{i,j}^{n,m})& \\
= \Delta t \Delta x \sum_{j=1}^{r} \sum_{m=1}^{s} \tilde{b}_m b_j (\mathrm{d}U_{i,j}^{n,m} \wedge (-V''(U_{i,j}^{n,m}) \mathrm{d}U_{i,j}^{n,m}))& \\
= 0.&
\end{aligned} \tag{2.16}$$

Thus, the discrete multisymplectic conservation law in Eq. (2.12) is satisfied. □

More generally, any multi-Hamiltonian PDE discretised in time and space by symplectic RK methods satisfies a discrete multisymplectic conservation law similar to Eq. (2.12). This will be seen in Section 3.2 as a special case of Theorem 3.2.1.

It can be noted from the proof of Theorem 2.1.1 (and later from Theorem 3.2.1) that the equations that one obtains by discretising a multi-Hamiltonian PDE in time and space by symplectic RK or PRK methods satisfy a discrete multisymplectic conservation law regardless of whether the discrete equations have unique (or any) solutions. The existence of a discrete multisymplectic conservation law satisfied by the discrete equations is simply a consequence of formal algebraic manipulation of the discrete equations.

In the past, several authors [13, 38, 60, 63] have given discretisations of Eq. (1.6), which they have shown to formally satisfy a discrete multisymplectic conservation law.

However, simply having a discretisation of a multi-Hamiltonian PDE that satisfies a discrete multisymplectic conservation law is insufficient to define a multisymplectic integrator. The system of discretised equations must form a well-defined integrator. A multisymplectic discretisation of a multi-Hamiltonian PDE may fail to form a well-defined integrator when one or more of the following problems occur:

(i) there may be no obvious choice of dependent variables;

(ii) the discrete equations may not be well defined locally (i.e., there may not be one equation per dependent variable per cell);

(iii) the discrete equations may not be well defined globally (i.e., there may not be one equation per dependent variable across all spatial grid points when boundary conditions are imposed);

(iv) the discrete equations may not have a solution, may not have a unique solution, may not have isolated solutions, or may have spurious solutions that are not representative of the continuous PDE.

Difficulties due to these problems already occur for the most popular multisymplectic integrator, the Preissman box scheme. With periodic boundary conditions in one spatial dimension, the discrete equations typically only have solutions with an odd number of grid points, while with an even number of grid points they have no solution (nonlinear problems) or an infinite number of solutions (linear problems). With higher order GRK methods these problems are even worse, as will be shown in the following sections.

2.2 The spatial discretisation

As mentioned above, to form a multisymplectic integrator a system of equations must do more than simply satisfy a discrete multisymplectic conservation law. In order to be a multisymplectic integrator, the system of equations must first be a well-defined numerical integrator.

Now, Eqs. (2.8) and (2.9) are obtained by first applying an r-stage symplectic GRK discretisation in space to the continuous equations to obtain a system of semi-discrete equations. This is followed by applying an s-stage symplectic GRK discretisation in time to the semi-discrete equations to obtain a fully discrete system of equations.

For now, let us consider the system of equations obtained from only applying the GRK discretisation in space. They are:

$$
\begin{aligned}
U_{i,j} &= u_i + \Delta x \sum_{k=1}^{r} a_{jk} \partial_x U_{i,k}, \qquad \text{for } j = 1, \dots, r,\\
W_{i,j} &= w_i + \Delta x \sum_{k=1}^{r} a_{jk} \partial_x W_{i,k}, \qquad \text{for } j = 1, \dots, r,\\
u_{i+1} &= u_i + \Delta x \sum_{j=1}^{r} b_j \partial_x U_{i,j},\\
w_{i+1} &= w_i + \Delta x \sum_{j=1}^{r} b_j \partial_x W_{i,j},
\end{aligned}
\tag{2.17}
$$

where

$$\begin{aligned}\partial_t U_{i,j} &= V_{i,j}\\ \partial_x U_{i,j} &= W_{i,j},\\ \partial_x W_{i,j} &= (\partial_t V_{i,j} + V'(U_{i,j})).\end{aligned} \tag{2.18}$$

Typically, in a RK method, the above equations define a map from (u_i, w_i) to (u_{i+1}, w_{i+1}), where the stage variables $U_{i,j}$ and $W_{i,j}$ are determined from the first two lines of Eq. (2.17) and from Eq. (2.18). However, it is not possible to uniquely determine the value of $U_{i,j}$ and $W_{i,j}$ from these equations when only u_i and w_i are known beforehand since there are more unknown variables than there are equations, i.e., there are $5r + 2$ equations ($2r+2$ in Eq. (2.17) and $3r$ in Eq. (2.18)) and $7r+2$ unknowns (u_{i+1}, w_{i+1}, $U_{i,j}$, $W_{i,j}$, $V_{i,j}$, $\partial_x U_{i,j}$, $\partial_x W_{i,j}$, $\partial_t U_{i,j}$, and $\partial_t V_{i,j}$). So the usual interpretation of Eqs. (2.17) and (2.18) as a map from (u_i, w_i) to (u_{i+1}, w_{i+1}) is not well defined. (If the GRK discretisation is applied in time instead of space, then a set of equations similar to Eqs. (2.17) and (2.18) are obtained, which are likewise not well defined as a map from (u^n, v^n) to (u^{n+1}, v^{n+1}).)

Since Eqs. (2.8) and (2.9) are obtained by applying a symplectic GRK discretisation to Eqs. (2.17) and (2.18), one might expect the fully discrete system of equations to also fail to be well defined. This will be considered further in Section 2.3, but for now, a different interpretation of Eqs. (2.17) and (2.18) will be considered.

2.2.1 Implicit ODEs

In the previous section it was shown that the RK discretisation in space of the nonlinear wave equation fails to form a well-defined integrator based on the interpretation of the discretised equations as a map from (u_i, w_i) to (u_{i+1}, w_{i+1}). This was due to an imbalance in the number of equations and the number of unknown variables, in particular, the variables causing this imbalance are $V_{i,j}$, $\partial_t V_{i,j}$ and $\partial_t U_{i,j}$.

If, instead of a map from (u_i, w_i) to (u_{i+1}, w_{i+1}), Eqs (2.17) and (2.18) are interpreted as a system of implicit ODEs in time for the variables $U_{i,j}$ and $V_{i,j}$ (i.e., the variables u_i, w_i, $W_{i,j}$, $\partial_x U_{i,j}$, $\partial_x W_{i,j}$, $\partial_t U_{i,j}$ and $\partial_t V_{i,j}$ are all written in terms of $U_{i,j}$ and $V_{i,j}$), then the number of unknown variables per cell does match the number of equations per cell.

In general, for a multi-Hamiltonian PDE (1.6), $\boldsymbol{L}$ may be written in Darboux normal form and the PDE may then be written as

$$\boldsymbol{L}\boldsymbol{z}_x = \nabla_{\boldsymbol{z}} S(\boldsymbol{z}) - \boldsymbol{K}\boldsymbol{z}_t. \tag{2.19}$$

Discretising Eq. (2.19) with an r-stage RK method gives a system of implicit differential-algebraic equations (i.e., Eq. (1.81)) for the node and stage variables in the RK discretisation, where the stage variables satisfy

$$\boldsymbol{L}\partial_x \boldsymbol{Z}_{i,j} = \nabla_{\boldsymbol{Z}_{i,j}} S(\boldsymbol{Z}_{i,j}) - \boldsymbol{K}\partial_t \boldsymbol{Z}_{i,j}. \tag{2.20}$$

The node variables in this discretisation can be eliminated by applying a forward difference operator δ^+ across cells (i.e., $\delta^+ \boldsymbol{Z}_{i,j} = \boldsymbol{Z}_{i+1,j} - \boldsymbol{Z}_{i,j}$) and multiplying by $\boldsymbol{L}$ to get

$$\begin{aligned}\delta^+ \boldsymbol{L}\boldsymbol{Z}_{i,j} &= \delta^+ \boldsymbol{L}\boldsymbol{z}_i + \Delta x \sum_{k=1}^{r} a_{jk}\delta^+ \boldsymbol{L}\partial_x \boldsymbol{Z}_{i,k} \\ &= \Delta x \sum_{k=1}^{r} (b_k + a_{jk}\delta^+)\boldsymbol{L}\partial_x \boldsymbol{Z}_{i,k}.\end{aligned} \tag{2.21}$$

Substituting the variables $\boldsymbol{L}\partial_x \boldsymbol{Z}_{i,j}$ from Eq. (2.20) into Eq. (2.21) gives

$$\delta^+ \boldsymbol{L}\boldsymbol{Z}_{i,j} = \Delta x \sum_{k=1}^{r} (b_k + a_{jk}\delta^+)(\nabla_{\boldsymbol{Z}_{i,k}} S(\boldsymbol{Z}_{i,k}) - \boldsymbol{K}\partial_t \boldsymbol{Z}_{i,k}) \tag{2.22}$$

or, in tensor form

$$\delta^+ \boldsymbol{L}\boldsymbol{Z}_i = \Delta x(\boldsymbol{1}\boldsymbol{b}^T + \boldsymbol{A}\delta^+)(\nabla_{\boldsymbol{Z}_i} S(\boldsymbol{Z}_i) - \boldsymbol{K}\partial_t \boldsymbol{Z}_i), \tag{2.23}$$

where $\boldsymbol{Z}_i$ is the tensor of the stage variables of $\boldsymbol{z}$ in cell i, $\boldsymbol{1}$ is a vector with all entries set to 1, $\boldsymbol{b}$ is the row vector of b_k values and $\boldsymbol{A}$ is the matrix of a_{jk} values.

These are implicit ODEs in time for the $\boldsymbol{Z}_i$ variables. If the operator $(\boldsymbol{1}\boldsymbol{b}^T + \boldsymbol{A}\delta^+)$ is non-singular then it may be formally inverted to write

$$\boldsymbol{K}\partial_t \boldsymbol{Z}_i + \frac{1}{\Delta x}(\boldsymbol{1}\boldsymbol{b}^T + \boldsymbol{A}\delta^+)^{-1}\delta^+ \boldsymbol{L}\boldsymbol{Z}_i = \nabla_{\boldsymbol{Z}_i} S(\boldsymbol{Z}_i). \tag{2.24}$$

Thus the effect of the RK discretisation in space is to approximate the operator $\frac{\partial}{\partial x}$ by the finite difference operator $\frac{1}{\Delta x}(\boldsymbol{1}\boldsymbol{b}^T + \boldsymbol{A}\delta^+)^{-1}\delta^+$, which does not depend on the PDE being discretised (although it is, in general, implicit). Since this approximation to $\frac{\partial}{\partial x}$ is linear, most of its properties can be understood from the behaviour of the RK method applied to linear differential equations, which is an extremely well-understood subject.

If $\boldsymbol{L}$ is singular in Eq. (2.19), then some of the components of the PDE will contain no spatial derivatives and it is questionable as to whether these components should be discretised in space. However, it can be seen from Eq. (2.24) that the terms that vanish in the PDE also vanish in the semi-discretisation, so no harm is done. Similarly, if $\boldsymbol{K}$ is singular, then some of the components of the PDE correspond to constraints on the $\boldsymbol{Z}_i$. In applications, these constraints will be eliminated in both the PDE and the semi-discretisation, leading to $r \times \mathrm{rank}(\boldsymbol{K})$ equations per cell. In other words, the RK discretisation in space does not affect the structure (i.e., $\boldsymbol{K}$, $\boldsymbol{L}$, and $S(\boldsymbol{z})$) of the multi-Hamiltonian PDE.

While this interpretation of the equations that one obtains by discretising a multi-Hamiltonian PDE in space with a RK method as a system of implicit ODEs may not be sufficient to give a well-defined multisymplectic integrator, since the operator $(\boldsymbol{1}\boldsymbol{b}^T + \boldsymbol{A}\delta^+)$ may be singular, it does, in general, avoid problems (i) and (ii) of Section 2.1.

2.2.2 Odd behaviour

As with many other approaches to forming numerical integrators, solving the system of equations given by Eqs. (2.17) and (2.18) for u_i, w_i, $W_{i,j}$, $\partial_x U_{i,j}$, $\partial_x W_{i,j}$, $\partial_t U_{i,j}$ and $\partial_t V_{i,j}$ requires solving a system of linear equations. Moreover, since Eq. (2.17) is obtained from a Gaussian RK discretisation and couples together the variables from neighbouring cells (through lines 3 and 4 of Eq. (2.17)), it is necessary to solve for all cells simultaneously.

For the nonlinear wave equation with periodic boundary conditions, we can substitute Eq. (2.18) into Eq. (2.17) and then write the first and third lines of Eq. (2.17) as the following matrix equation:

$$\left[\begin{array}{ccc|cccc}
\Delta x\boldsymbol{A} & & & \boldsymbol{1} & & & \\
& \ddots & & & \ddots & & \\
& & \ddots & & & \ddots & \\
& & \Delta x\boldsymbol{A} & & & & \boldsymbol{1} \\
\hline
\Delta x\boldsymbol{b} & & & 1 & -1 & & \\
& \ddots & & & \ddots & \ddots & \\
& & \ddots & & & \ddots & -1 \\
& & \Delta x\boldsymbol{b} & -1 & & & 1
\end{array}\right]
\left[\begin{array}{c}
\boldsymbol{W}_1 \\ \vdots \\ \vdots \\ \boldsymbol{W}_N \\ \hline u_1 \\ \vdots \\ \vdots \\ u_N
\end{array}\right]
=
\left[\begin{array}{c}
\boldsymbol{U}_1 \\ \vdots \\ \vdots \\ \boldsymbol{U}_N \\ \hline 0 \\ \vdots \\ \vdots \\ 0
\end{array}\right], \tag{2.25}$$

where $\boldsymbol{A}$ is the $r \times r$ matrix of a_{ij} values in the GRK method, $\boldsymbol{b}$ is the $1 \times r$ vector of b_j values, $\boldsymbol{1}$ is an $r \times 1$ vector with all entries set to 1, $\boldsymbol{W}_i$ and $\boldsymbol{U}_i$ are the vectors of stage variables of w and u in cell i, and all other entries in the large matrix are zero. Solving Eq. (2.25) for the variables $\boldsymbol{W}_i$ and u_i requires finding the inverse of the $N(r+1) \times N(r+1)$ matrix.

In general, for a multi-Hamiltonian PDE written in the form of Eq. (2.19) with periodic boundary conditions and discretised in space with an r-stage GRK discretisation, the matrix equation equivalent to Eq. (2.25) is given by

$$\left[\begin{array}{ccc|cccc}
\Delta x\boldsymbol{A} & & & \boldsymbol{1} & & & \\
& \ddots & & & \ddots & & \\
& & \ddots & & & \ddots & \\
& & \Delta x\boldsymbol{A} & & & & \boldsymbol{1} \\
\hline
\Delta x\boldsymbol{b} & & & \boldsymbol{e} & -\boldsymbol{e} & & \\
& \ddots & & & \ddots & \ddots & \\
& & \ddots & & & \ddots & -\boldsymbol{e} \\
& & \Delta x\boldsymbol{b} & -\boldsymbol{e} & & & \boldsymbol{e}
\end{array}\right]
\left[\begin{array}{c}
(\nabla_{\boldsymbol{z}} S(\boldsymbol{Z}) - \boldsymbol{K}\partial_t \boldsymbol{Z})_1 \\ \vdots \\ \vdots \\ (\nabla_{\boldsymbol{z}} S(\boldsymbol{Z}) - \boldsymbol{K}\partial_t \boldsymbol{Z})_N \\ \hline \boldsymbol{L}\boldsymbol{z}_1 \\ \vdots \\ \vdots \\ \boldsymbol{L}\boldsymbol{z}_N
\end{array}\right]
=
\left[\begin{array}{c}
\boldsymbol{L}\boldsymbol{Z}_1 \\ \vdots \\ \vdots \\ \boldsymbol{L}\boldsymbol{Z}_N \\ \hline 0 \\ \vdots \\ \vdots \\ 0
\end{array}\right], \tag{2.26}$$

where $\boldsymbol{e} = [1,\ 0,\ \ldots\ 0]$ and $(\nabla_{\boldsymbol{z}} S(\boldsymbol{Z}) - \boldsymbol{K}\partial_t \boldsymbol{Z})_i$ is the vector $\nabla_{\boldsymbol{z}} S(\boldsymbol{Z}) - \boldsymbol{K}\partial_t \boldsymbol{Z}$ in cell i.

The size of the matrix on the left is now $Nn(r+1) \times Nn(r+1)$.

For a general multi-Hamiltonian PDE, not only does the matrix on the left of Eq. (2.26) need to be invertible (inverting this matrix is equivalent to finding the operator $\frac{1}{\Delta x}(\mathbf{1}\boldsymbol{b}^T + \boldsymbol{A}\delta^+)^{-1}\delta^+$), but the variables in the system that do not have time derivatives need to be eliminated so as to write Eq. (2.26) as a system of (either explicit or implicit) ODEs in terms of the variables that do have time derivatives. However, for systems in which the variables without time derivatives appear linearly in $\nabla_{\boldsymbol{z}} S(\boldsymbol{z})$, this extra requirement is always satisfied. The conditions such that the matrix on the left of Eq. (2.26) is invertible are given in the following theorem.

Theorem 2.2.1. *The matrix on the left-hand side of Eq. (2.26) is only non-singular when both the stage order, r, and the number of grid points, N, are odd.*

Proof. The determinant of a block matrix,

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right], \tag{2.27}$$

is given by $\det(A)\det(D - CA^{-1}B)$.

Since Eq. (2.19) is discretised by a Gaussian RK method, the matrix $\boldsymbol{A}$ in Eq. (2.26) is of full rank, therefore the upper left block of the matrix in Eq. (2.26) has determinant $(\Delta x)^{rNn}\det(\boldsymbol{A})^{Nn}$. Also, the block matrix product $CA^{-1}B$, where A, B and C are the blocks in Eq. (2.26) matching those of Eq. (2.27), is the $Nn \times Nn$ identity matrix scaled by $\boldsymbol{b}\boldsymbol{A}^{-1}\mathbf{1}$.

Now, the stability function for a RK method is given as [19]

$$R(z) = 1 + z\boldsymbol{b}^T(\boldsymbol{I} - z\boldsymbol{A})^{-1}\mathbf{1}, \tag{2.28}$$

which, in the limit $z \to \infty$, becomes

$$\lim_{z\to\infty} R(z) = 1 - \boldsymbol{b}\boldsymbol{A}^{-1}\mathbf{1}. \tag{2.29}$$

On the other hand, it has been shown [24] that for an r-stage Gaussian RK method, the stability function is the diagonal Padé approximation

$$R(z) = P_{rr}(z) = \frac{N_{rr}(z)}{N_{rr}(-z)}, \tag{2.30}$$

which approaches $(-1)^r$ as $z \to \infty$. Comparing these two limits gives $1 - \boldsymbol{b}\boldsymbol{A}^{-1}\mathbf{1} = (-1)^r$, and thus we obtain

$$\boldsymbol{b}\boldsymbol{A}^{-1}\mathbf{1} = \begin{cases} 0 & \text{if } r \text{ is even,} \\ 2 & \text{if } r \text{ is odd.} \end{cases} \tag{2.31}$$

Hence if r is even, the block matrix product, $D - CA^{-1}B$, is the $Nn \times Nn$ matrix

$$\begin{bmatrix} \boldsymbol{e} & -\boldsymbol{e} & & \\ & \ddots & \ddots & \\ & & \ddots & -\boldsymbol{e} \\ -\boldsymbol{e} & & & \boldsymbol{e} \end{bmatrix}, \tag{2.32}$$

which is always singular. But when r is odd, this block matrix product is the $Nn \times Nn$ matrix

$$\begin{bmatrix} -\boldsymbol{e} & -\boldsymbol{e} & & \\ & \ddots & \ddots & \\ & & \ddots & -\boldsymbol{e} \\ -\boldsymbol{e} & & & -\boldsymbol{e} \end{bmatrix}, \tag{2.33}$$

which is singular when N is even and non-singular (with determinant $(-2)^n$) when N is odd.

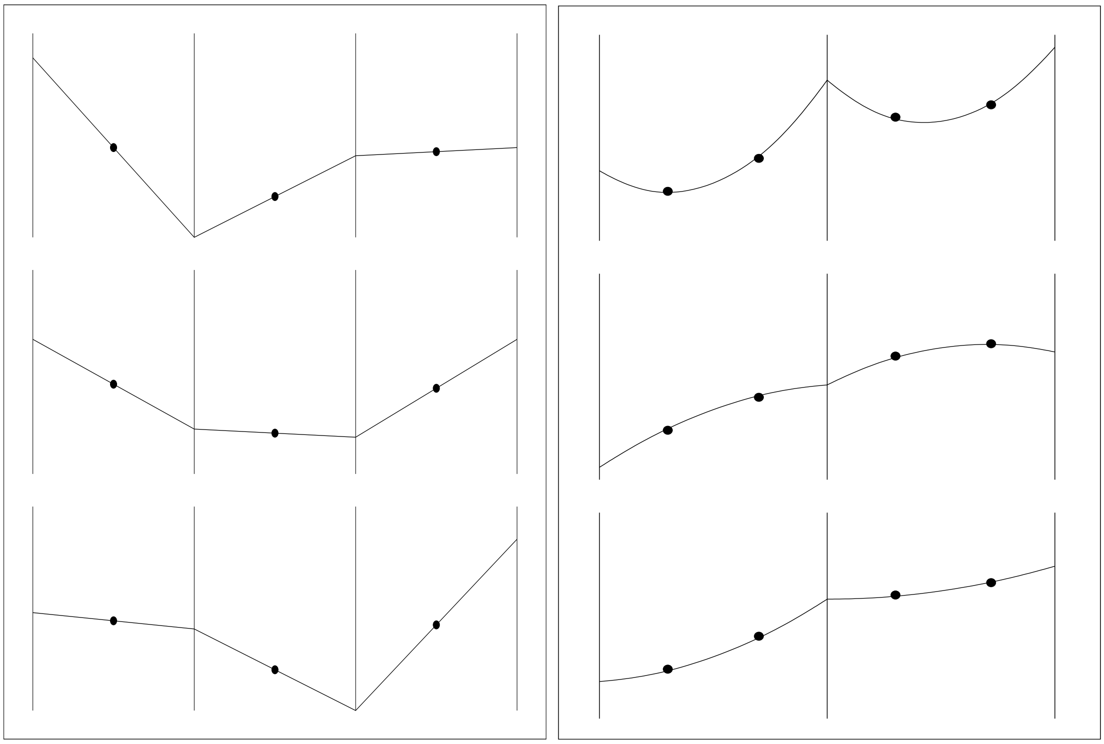
Therefore, the matrix on the left-hand side of Eq. (2.26) is only non-singular (with determinant $(-2)^n(\Delta x)^{rNn}\det(\boldsymbol{A})^{Nn}$) when both r and N are odd. □

This difficulty has been partially recognised before; the commonly used implicit mid-point method (GRK with $r = 1$) has long been known to only work for an odd number of grid points. A pictorial representation of why Theorem 2.2.1 holds may be seen in Figure 2.2.

From Figure 2.2 we can see that Theorem 2.2.1 should also hold true for a GRK discretisation in space on a non-uniform grid as the width of each cell does not affect the amount the collocation polynomial on the right boundary of that cell changes for a given change on the left boundary of that cell. Indeed, this can also be seen from the above proof of Theorem 2.2.1 since the block A in Eq. (2.27) is block-diagonal and $D - CA^{-1}B$ is independent of Δx. The only modification to the determinant resulting from the non-uniform grid is the $(\Delta x)^{rNn}$ factor becomes $\prod_{i=1}^{N}(\Delta x_i)^{rn}$, where Δx_i is the width of cell i.

Theorem 2.2.1 is stated for the case of periodic boundary conditions and it is not obvious from the above proof that this theorem should hold true for other types of boundary conditions. Similarly, Figure 2.2 is designed to illustrate Theorem 2.2.1 for the case of periodic boundary conditions, however, it also indicates the possibilities for other types of boundary conditions.

Dirichlet and Neumann boundary conditions are each defined by specifying the values of some of the entries in $\boldsymbol{z}$ on the boundaries. For Dirichlet boundary conditions it is the variables in the original PDE that are specified, e.g., u for the nonlinear wave equation,



(a) Odd number of cells and internal stages. (b) Even number of cells and internal stages.

Figure 2.2: Collocation polynomials for periodic boundary conditions. (a) As the value on the left boundary decreases, the value on the right boundary increases. At some value they match. (b) As the value on the left boundary decreases, the value on the right boundary decreases at the same rate.

whereas, for Neumann boundary conditions it is the variables corresponding to the spatial derivatives of the original variables, e.g., w for the nonlinear wave equation. Other boundary conditions (such as non-reflecting boundaries) may be defined which involve combinations of Dirichlet or Neumann boundary conditions.

Now, if the value of a variable is specified on the left boundary and all the stage variables are evolved forwards in time according to the PDE (i.e., the matrix corresponding to the one in Eq. (2.26) has full rank), then it can be seen from the collocation polynomials that the value of the variables on the right boundary will not, in general, satisfy a predefined boundary condition. However, it can be seen from the collocation polynomials in Figure 2.2 that for each of these types of boundary conditions, it is possible to satisfy a predefined right boundary condition (regardless of whether r and N are odd) by allowing one of the stage variables to be written as a function of the other stage variables instead of evolving according to the PDE. This corresponds to the matrix having nullity 1 in all the cases where it is singular.

If boundary conditions are not specified on the right boundary, then the above arguments will not apply, however, without boundary conditions on the right boundary, the discretisation may no longer be an accurate representation of the PDE; this will be discussed in Section 2.3.1 in the context of a full discretisation.

2.3 The full discretisation

It is possible to consider whether the full discretisation (Eqs. (2.8) and (2.9)) forms a well-defined numerical integrator irrespective of it being derived from a spatial discretisation that may not be well defined.

The system of equations that will be considered throughout this section are those obtained by substituting Eq. (2.9) into Eq. (2.8). They are

$$\begin{aligned}
U_{i,j}^{n,m} &= u_i^{n,m} + \Delta x \sum_{k=1}^{r} a_{jk} W_{i,k}^{n,m},\\
W_{i,j}^{n,m} &= w_i^{n,m} + \Delta x \sum_{k=1}^{r} a_{jk} (V'(U_{i,j}^{n,m}) + \partial_t V_{i,j}^{n,m}),\\
U_{i,j}^{n,m} &= u_{i,j}^{n} + \Delta t \sum_{l=1}^{s} \tilde{a}_{ml} V_{i,j}^{n,l},\\
V_{i,j}^{n,m} &= v_{i,j}^{n} + \Delta t \sum_{l=1}^{s} \tilde{a}_{ml} \partial_t V_{i,j}^{n,l},\\
u_{i+1}^{n,m} &= u_i^{n,m} + \Delta x \sum_{k=1}^{r} b_k W_{i,k}^{n,m},\\
w_{i+1}^{n,m} &= w_i^{n,m} + \Delta x \sum_{k=1}^{r} b_k (V'(U_{i,j}^{n,m}) + \partial_t V_{i,j}^{n,m}),\\
u_{i,j}^{n+1} &= u_{i,j}^{n} + \Delta t \sum_{l=1}^{s} \tilde{b}_l V_{i,j}^{n,l},\\
v_{i,j}^{n+1} &= v_{i,j}^{n} + \Delta t \sum_{l=1}^{s} \tilde{b}_l \partial_t V_{i,j}^{n,l},
\end{aligned} \tag{2.34}$$

where $1 \leq j \leq r$, $1 \leq m \leq s$.

This system of equations couples $4rs + 4r + 4s$ variables ($U_{i,j}^{n,m}$, $W_{i,j}^{n,m}$, $V_{i,j}^{n,m}$, $\partial_t V_{i,j}^{n,m}$, $u_{i,j}^{n}$, $u_{i,j}^{n+1}$, $u_i^{n,m}$, $u_{i+1}^{n,m}$, $w_i^{n,m}$, $w_{i+1}^{n,m}$, $v_{i,j}^{n}$, $v_{i,j}^{n+1}$) through $4rs + 2r + 2s$ equations per cell. Therefore, in order to have a chance of being well defined, it is necessary that precisely $2r + 2s$ variables be given as boundary conditions. Since we want to have an integrator that steps forward in time, the $2r$ variables that will be chosen as boundary conditions will be the variables $u_{i,j}^{n}$ and $v_{i,j}^{n}$ for some value of n (usually $n = 0$). The $2s$ variables can be chosen in several different ways and in the following sections I will consider whether the integrator formed by these choices of boundary conditions give rise to well-defined numerical integrators and, if so, whether the integrator is practical.

An important point to be considered here is that the choice of boundary conditions will, in general, break the space-time symmetry that is inherent in multisymplectic theory. That is, in a multi-Hamiltonian PDE (and in the discretised system of equations), the time and spatial dimensions are treated on an equal footing, however, when boundary conditions are applied to the PDE (or the discrete equations), the boundary conditions in space are, in general, treated differently from the boundary conditions in time.

2.3.1 Local integrators

The first choice of variables to specify the $2s$ boundary conditions on are the variables $u_0^{n,m}$ and $w_0^{n,m}$, i.e., on the left boundary. Specifying the boundary conditions on these variables allows one to form an integrator that can be solved locally (i.e., within a single cell) to provide the value of the variables $u_{i+1}^{n,m}$ and $w_{i+1}^{n,m}$ for the cell to the right and the value of the variables $u_{i,j}^{n+1}$ and $v_{i,j}^{n+1}$ for the cell above. Thus allowing the solution of the integrator to propagate forwards throughout time and space from a lower boundary in time and a left boundary in space.

Since the matrix formed by the $\tilde{a}_{ij}$ coefficients of the GRK discretisation in time ($\tilde{\boldsymbol{A}}$) has full rank, the first four lines of Eq. (2.34) can be rearranged as follows:

$$\begin{aligned}
\partial_t V_{i,j}^{n,m} &= \frac{1}{\Delta t}\sum_{l=1}^{s}(\tilde{\boldsymbol{A}}^{-1})_{ml}(V_{i,j}^{n,l} - v_{i,j}^{n}),\\
W_{i,j}^{n,m} &= w_i^{n,m} + \Delta x\sum_{k=1}^{r} a_{jk}V'(U_{i,k}^{n,m}) + \frac{\Delta x}{\Delta t}\sum_{k=1}^{r} a_{jk}\sum_{l=1}^{s}(\tilde{\boldsymbol{A}}^{-1})_{ml}(V_{i,k}^{n,l} - v_{i,k}^{n}),\\
V_{i,j}^{n,m} &= \frac{1}{\Delta t}\sum_{l=1}^{s}(\tilde{\boldsymbol{A}}^{-1})_{ml}(U_{i,j}^{n,l} - u_{i,j}^{n}),\\
U_{i,j}^{n,m} &= u_i^{n,m} + \Delta x\sum_{k=1}^{r} a_{jk}\left(w_i^{n,m} + \Delta x\sum_{p=1}^{r} a_{kp}V'(U_{i,p}^{n,m})\right.\\
&\qquad\left. + \frac{\Delta x}{\Delta t}\sum_{p=1}^{r} a_{kp}\sum_{l=1}^{s}(\tilde{\boldsymbol{A}}^{-1})_{ml}\left(\frac{1}{\Delta t}\sum_{q=1}^{s}(\tilde{\boldsymbol{A}}^{-1})_{lq}(U_{i,p}^{n,q} - u_{i,p}^{n}) - v_{i,p}^{n}\right)\right).
\end{aligned} \tag{2.35}$$

These equations may be solved (in the reverse order as presented), with the equation for $U_{i,j}^{n,m}$ being implicit while the rest are explicit. The values of $U_{i,j}^{n,m}$, $W_{i,j}^{n,m}$, $V_{i,j}^{n,m}$ and $\partial_t V_{i,j}^{n,m}$ may then be substituted into the last four lines of Eq. (2.34) to obtain the values of $u_{i+1}^{n,m}$, $w_{i+1}^{n,m}$, $u_{i,j}^{n+1}$ and $v_{i,j}^{n+1}$.

Now, in order to calculate the value of the variables at the corners of a cell (i.e., at the nodes of the grid) a further set of equations are required as Eq. (2.34) only gives the value of the variables along the edge of each cell. These extra equations are given by

$$
\begin{aligned}
u_{i,j}^n &= u_i^n + \Delta x \sum_{k=1}^{r} a_{jk} w_{i,k}^n, & u_{i+1}^n &= u_i^n + \Delta x \sum_{k=1}^{r} b_k w_{i,k}^n, \\
w_{i,j}^n &= w_i^n + \Delta x \sum_{k=1}^{r} a_{jk} (V'(u_{i,k}^n) + \partial_t v_{i,k}^n), & w_{i+1}^n &= w_i^n + \Delta x \sum_{k=1}^{r} b_k (V'(u_{i,k}^n) + \partial_t v_{i,k}^n), \\
u_i^{n,m} &= u_i^n + \Delta t \sum_{l=1}^{s} \tilde{a}_{ml} v_i^{n,l}, & u_i^{n+1} &= u_i^n + \Delta t \sum_{l=1}^{s} \tilde{b}_l v_i^{n,l}, \\
v_i^{n,m} &= v_i^n + \Delta t \sum_{l=1}^{s} \tilde{a}_{ml} \partial_t v_i^{n,l}, & v_i^{n+1} &= v_i^n + \Delta t \sum_{l=1}^{s} \tilde{b}_l \partial_t v_i^{n,l},
\end{aligned}
\tag{2.36}
$$

where $u_i^{n,m}$ and $u_{i,j}^n$ are known from Eq. (2.34) and extra boundary conditions need to be specified for u_i^n, w_i^n and v_i^n. These equations are also required if one wishes to calculate the value of various properties of the solution. For example, the energy density and energy flux functions, which make up the energy conservation law, depend on the $w_{i,j}^n$ and $v_i^{n,m}$ variables.

Since the matrices of the $\tilde{a}_{ij}$ and a_{ij} coefficients in the GRK methods in time and space are of full rank, we can calculate the value of $w_{i,j}^n$, $v_i^{n,m}$, $\partial_t v_{i,j}^n$ and $\partial_t v_i^{n,m}$ from the equations in the left column of Eq. (2.36). The value of these variables can then be substituted into the right column of Eq. (2.36) to get u_{i+1}^n, w_{i+1}^n, u_i^{n+1} and v_i^{n+1}.

This method seems promising at first as it can solve for each cell based on the values of the neighbouring cells. However, on closer inspection it is revealed that the method contains some critical flaws. Specifically,

- u_i^n for a given cell can be found from the cell to the left by the first line of the right column of Eq. (2.36) or from the cell below by the third line of the right column of Eq. (2.36) and there is no guarantee that these equations are consistent, so while Eq. (2.36) is locally well defined within a cell, it is overdetermined when coupled to neighbouring cells. This is an example of problem (iii) in Section 2.1.

- More importantly, the Cauchy data for the numerical integrator formed by Eq. (2.34) and this choice of boundary conditions doubly intersects the characteristics of the PDE.

  That is, one solution to the wave equation is a travelling wave that is moving to the left, which is represented in phase space by a characteristic with negative slope. This characteristic crosses both $t = 0$ and $x = 0$, where we have specified both Dirichlet and Neumann boundary conditions. This abundance of boundary conditions means that this solution is overdetermined and, in general, will not exist.

  Thus, the solution obtained from the numerical integrator does not represent a solution of the continuous PDE. This is an example of problem (iv) in Section 2.1.

The other choice of the $2s$ variables to be used to specify the boundary conditions while retaining local solvability are the variables $u_N^{n,m}$ and $v_N^{n,m}$, where N is the number of cells in space, i.e., both on the right boundary. The only difference between this choice and the previous one is that the solution propagates left from a right boundary in space. Hence, the integrator formed from this choice suffers the same flaws as the previous integrator.

2.3.2 Global integrators

Another choice of variables to specify the $2s$ boundary conditions on are the variables $u_0^{n,m}$ and $u_N^{n,m}$, i.e., one on each of the left and right boundaries. Since the boundary conditions are not local to a single cell, the integrator formed by this choice of boundary conditions needs to be solved globally in space to advance from one time level to the next.

The integrator formed by this choice of boundary conditions may be written as

$$\begin{aligned}
U_{i,j}^{n,m} &= u_0^{n,m} + \Delta x \sum_{k=1}^{r} \left(b_k \sum_{p=1}^{i-1} W_{p,k}^{n,m} + a_{jk} W_{i,k}^{n,m} \right), \\
W_{i,j}^{n,m} &= w_0^{n,m} + \Delta x \sum_{k=1}^{r} \left(b_k \sum_{p=1}^{i-1} \left(\partial_t V_{p,k}^{n,m} + V'(U_{p,k}^{n,m}) \right) + a_{jk} \left(\partial_t V_{i,k}^{n,m} + V'(U_{i,k}^{n,m}) \right) \right), \\
U_{i,j}^{n,m} &= u_{i,j}^{n} + \Delta t \sum_{l=1}^{s} \tilde{a}_{ml} V_{i,j}^{n,l}, \\
V_{i,j}^{n,m} &= v_{i,j}^{n} + \Delta t \sum_{l=1}^{s} \tilde{a}_{ml} \partial_t V_{i,j}^{n,l}, \\
u_N^{n,m} &= u_0^{n,m} + \Delta x \sum_{k=1}^{r} \left(b_k \sum_{p=1}^{N-1} W_{p,k}^{n,m} \right), \\
w_N^{n,m} &= w_0^{n,m} + \Delta x \sum_{k=1}^{r} \left(b_k \sum_{p=1}^{N-1} \left(\partial_t V_{p,k}^{n,m} + V'(U_{p,k}^{n,m}) \right) \right), \\
u_{i,j}^{n+1} &= u_{i,j}^{n} + \Delta t \sum_{l=1}^{s} \tilde{b}_l V_{i,j}^{n,l}, \\
v_{i,j}^{n+1} &= v_{i,j}^{n} + \Delta t \sum_{l=1}^{s} \tilde{b}_l \partial_t V_{i,j}^{n,l}.
\end{aligned} \tag{2.37}$$

The first five lines of Eq. (2.37) can be written in the form of a matrix equation similar to Eq. (2.25), but with the variables $u_0^{n,m}$, $u_{i,j}^n$, $v_{i,j}^n$, and $(u_N^{n,m} - u_0^{n,m})$ on the right and the variables $U_{i,j}^{n,m}$, $V_{i,j}^{n,m}$, $W_{i,j}^{n,m}$, $\partial_t V_{i,j}^{n,m}$, and $w_0^{n,m}$ on the left being multiplied by a $(4Nrs+s)\times(4Nrs+s)$ matrix. Using this integrator requires solving this matrix equation for $U_{i,j}^{n,m}$, $V_{i,j}^{n,m}$, $W_{i,j}^{n,m}$, $\partial_t V_{i,j}^{n,m}$ and $w_0^{n,m}$ then using the last two lines of Eq. (2.37) to advance the values of $u_{i,j}^n$ and $v_{i,j}^n$ to the next time level. The sixth line of Eq. (2.37) may be used to find $w_N^{n,m}$, although this is not required to step the system forwards in time.

Unlike the matrix on the left of Eq. (2.25), this $(4Nrs + s) \times (4Nrs + s)$ matrix requires neither r, s, nor N to be odd in order to be non-singular. However, when the term $V'(u)$ in the PDE is nonlinear, the matrix equation is also nonlinear and must be solved implicitly. In contrast to the local integrator in the previous section, this global integrator does appear to be well defined in the sense that the implicit matrix equation can be solved and the characteristics of the PDE meet the appropriate amount of Cauchy data at the boundaries. Nevertheless, this approach is impractical for the following two reasons.

Firstly, since we have to solve for an entire time level simultaneously, this large implicit matrix equation must be solved repeatedly for each step forwards in time. For any reasonable simulation, this will be a very expensive exercise in terms of data storage and computational time.

Secondly, this large matrix is very poorly conditioned and rapidly becomes worse as the order of the GRK discretisations in time and space are increased. Thus, a significant amount of error is introduced at each time step that may be much larger than the error expected from a high order GRK discretisation.

Similar choices of the $2s$ variables on which the boundary conditions are to be specified ($w_0^{n,m}$ and $w_N^{n,m}$, $u_0^{n,m}$ and $w_N^{n,m}$, or $w_0^{n,m}$ and $u_N^{n,m}$) result in the same set of equations as Eq. (2.37). While the implicit matrix equation is modified slightly depending on which variables are known and which are to be solved for, each of these choices of boundary conditions is impractical for the same reasons as mentioned above.

The last choice of variables for the $2s$ boundary conditions is to avoid specifying the value on the boundary, but rather to define periodic boundary conditions in space, i.e., $u_0^{n,m} = u_N^{n,m}$ and $w_0^{n,m} = w_N^{n,m}$. Once again, this results in an implicit matrix equation that is similar to the previous ones and is impractical for the same reasons.

2.3.3 The box scheme

The box scheme (sometimes known as the Preissman box scheme) is a concatenation of the implicit midpoint method in both time and space and, for simplicity, is often implemented with periodic boundary conditions. The implicit midpoint method is a symplectic Runge–Kutta method with $r = 1$ and, when suitable boundary conditions are chosen, is an example of when concatenating symplectic Runge–Kutta methods *does* result in a practical, well defined, multisymplectic integrator.

For the boundary conditions in Section 2.3.2, the box scheme requires inverting only a $(Nn+1) \times (Nn+1)$ matrix since the implicit midpoint method has only 1 internal stage. This matrix has a much simpler structure than the matrix obtained from a higher order discretisation in time or space and, for reasonable values of N, is not too badly conditioned. Moreover, the stage variables can be eliminated leaving just the node variables (this cannot be done for RK methods with more than 1 internal stage), thus the box scheme typically avoids the problems that plagued the methods in Section 2.3.2. These problems may be further reduced by eliminating most of the variables in the multi-Hamiltonian PDE and writing the PDE in terms of as few variables as possible, as is demonstrated in the KdV example below.

The usual method of constructing a box scheme is to first apply the following finite difference and average operators (which are equivalent to applying the implicit midpoint method then eliminating the stage variables) to the multi-Hamiltonian PDE (1.6):

$$\begin{aligned} D_t u_{i,j}^n &= \frac{1}{\Delta t}(u_{i,j}^{n+1} - u_{i,j}^n) = \frac{1}{\Delta t}[\ -1 \quad 1\ ]u_{i,j}^n, \\ M_t u_{i,j}^n &= \frac{1}{2}(u_{i,j}^{n+1} + u_{i,j}^n) = \frac{1}{2}[\ 1 \quad 1\ ]u_{i,j}^n, \end{aligned} \tag{2.38}$$

where stencil notation has been used. The result is

$$\frac{1}{2\Delta t}\boldsymbol{K}\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}\boldsymbol{z} + \frac{1}{2\Delta x}\boldsymbol{L}\begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}\boldsymbol{z} = \nabla_{\boldsymbol{z}} S\left(\frac{1}{4}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\boldsymbol{z}\right). \tag{2.39}$$

Next, the variables that were introduced to write the PDE as a multi-Hamiltonian PDE are eliminated to get a stencil in terms of the variables in the original PDE (typically, only a single variable remains). For example, the box scheme applied to the KdV equation,

$$u_t = V'(u)_x + \nu u_{xxx}, \tag{2.40}$$

results in the 12-pt stencil [80]

$$\begin{aligned} \frac{1}{16\Delta t}\begin{bmatrix} 1 & 3 & 3 & 1 \\ 0 & 0 & 0 & 0 \\ -1 & -3 & -3 & -1 \end{bmatrix} u = \frac{1}{4\Delta x}&\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} V'\left(\frac{1}{4}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} u\right) \\ &+ \frac{\nu}{4(\Delta x)^3}\begin{bmatrix} -1 & 3 & -3 & 1 \\ -2 & 6 & -6 & 2 \\ -1 & 3 & -3 & 1 \end{bmatrix} u. \end{aligned} \tag{2.41}$$

2.3.4 The box scheme in higher dimensions

The generalisation of the box scheme to higher dimensions is to simply apply the implicit midpoint method in each dimension.

As an example, consider the nonlinear wave equation in two spatial dimensions,

$$u_{tt} - u_{xx} - u_{yy} = V'(u). \tag{2.42}$$

This PDE has the following multi-Hamiltonian form when written as a first order system,

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \boldsymbol{z}_t + \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \boldsymbol{z}_x + \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \boldsymbol{z}_y = \nabla_{\boldsymbol{z}} S(\boldsymbol{z}) \tag{2.43}$$

where $\boldsymbol{z} = [u,\ v,\ w,\ \phi]^T$ and $S(z) = -V(u) + \frac{1}{2}(-v^2 + w^2 + \phi^2)$.

Applying the finite difference and average operators then eliminating the variables v, w and ϕ gives

$$D_t^2 M_x^2 M_y^2 u_{i,j}^n - D_x^2 M_t^2 M_y^2 u_{i,j}^n - D_y^2 M_t^2 M_x^2 u_{i,j}^n = M_t M_x M_y V'(M_t M_x M_y u_{i,j}^n). \tag{2.44}$$

Writing the x and y directions in stencil form gives

$$\begin{aligned} &\frac{1}{16\Delta t^2} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} (u_{i,j}^{n+2} - 2u_{i,j}^{n+1} + u_{i,j}^n) \\ &\qquad - \frac{1}{16\Delta x^2} \begin{bmatrix} 1 & -2 & 1 \\ 2 & -4 & 2 \\ 1 & -2 & 1 \end{bmatrix} (u_{i,j}^{n+2} + 2u_{i,j}^{n+1} + u_{i,j}^n) \\ &\qquad\qquad - \frac{1}{16\Delta y^2} \begin{bmatrix} 1 & 2 & 1 \\ -2 & -4 & -2 \\ 1 & 2 & 1 \end{bmatrix} (u_{i,j}^{n+2} + 2u_{i,j}^{n+1} + u_{i,j}^n) \\ &= \frac{1}{8} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \left(V' \left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \left(u_{i,j}^{n+2} + u_{i,j}^{n+1} \right) \right) + V' \left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \left(u_{i,j}^{n+1} + u_{i,j}^n \right) \right) \right). \end{aligned} \tag{2.45}$$

Chapter 3

Partitioned Runge–Kutta Discretisations

A brief description of two-partition partitioned Runge–Kutta (PRK) discretisations, which are typically used in the literature, was given in Section 1.4.3. In the following section I will give a more comprehensive description of a PRK discretisation. Following this I will show that, under a simple set of conditions, discretisation of a multi-Hamiltonian PDE in N time and spatial dimensions by independent PRK methods with an arbitrary number of partitions results in a system of equations that satisfy a discrete multisymplectic conservation law. This result both simplifies and generalises similar results on the multisymplecticity of PRK applied to multi-Hamiltonian PDEs that appear in the literature [38, 65].

The remainder of this chapter is concerned with the construction of explicit multisymplectic integrators based on the Lobatto IIIA–IIIB class of PRK discretisation. After introducing the Lobatto IIIA–IIIB class of methods, I will proceed to show how the discretisation in space of a class of multi-Hamiltonian PDEs by such methods leads to a system of explicit ODEs in time. This is done by way of a construction algorithm, which may be reduced to a simple two-step procedure. Several examples of multi-Hamiltonian PDEs are given to demonstrate how this construction algorithm may give explicit ODEs and also, how it may fail to do so. For each multi-Hamiltonian PDE that the construction algorithm fails to give a system of explicit ODEs, a 2-stage Lobatto IIIA–IIIB discretisation in space is applied and the equations are reduced to a single implicit ODE. Lastly, I will consider how these explicit ODEs handle periodic, Dirichlet and Neumann boundary conditions.

Much of the content of this chapter is contained in the paper "On multisymplecticity of partitioned Runge–Kutta methods" by myself and Robert McLachlan, which has been submitted to the SIAM Journal on Scientific Computing (SISC) [65].

3.1 Partitioned Runge–Kutta discretisation

When a non-autonomous ODE,

$$\boldsymbol{z}_t = f(t, \boldsymbol{z}(t)), \tag{3.1}$$

is discretised with a PRK discretisation the vector of dependent variables, $\boldsymbol{z} \in \mathbb{R}^n$, is partitioned into several partitions, $\boldsymbol{z}^{(\gamma)} \in \mathbb{R}^{n_\gamma}$, with $\sum_\gamma n_\gamma = n$. As before, a grid is then introduced where the grid points (or *nodes*) are taken (for convenience only) to have equal spacing Δt and the notation of Section 1.4.2 is extended as follows:

- let z^γ be the entry γ in $\boldsymbol{z}$,
- let $\boldsymbol{z}_i^{(\gamma)} \in \mathbb{R}^{n_\gamma}$ be the vector of variables in partition γ at the node in cell i,
- let $\boldsymbol{Z}_{i,j}^{(\gamma)} \in \mathbb{R}^{n_\gamma}$ be the vector of variables in partition γ at stage j in cell i,
- let the lack of a raised, parenthesised index, (γ), indicate the unpartitioned variable.

Then for an r-stage PRK discretisation of Eq. (3.1) one obtains a set of equations coupling the node values $\boldsymbol{z}_i$ to the stage values $\boldsymbol{Z}_{i,j}$ at r internal stages given by

$$\begin{aligned} \boldsymbol{Z}_{i,j}^{(\gamma)} &= \boldsymbol{z}_i^{(\gamma)} + \Delta t \sum_{k=1}^{r} a_{jk}^{(\gamma)} \partial_t \boldsymbol{Z}_{i,k}^{(\gamma)}, \quad j = 1, \ldots, r, \\ \boldsymbol{z}_{i+1}^{(\gamma)} &= \boldsymbol{z}_i^{(\gamma)} + \Delta t \sum_{j=1}^{r} b_j^{(\gamma)} \partial_t \boldsymbol{Z}_{i,j}^{(\gamma)}, \end{aligned} \tag{3.2}$$

for each γ, where the new variables, $\partial_t \boldsymbol{Z}_{i,j}$, satisfy Eq. (3.1), i.e.,

$$\partial_t \boldsymbol{Z}_{i,j} = f((i + c_j)\Delta t, \boldsymbol{Z}_{i,j}), \tag{3.3}$$

and the coefficients $b_j^{(\gamma)}$ and $a_{jk}^{(\gamma)}$ are chosen to satisfy certain order conditions.

When the number of partitions in the PRK method is one, i.e., $\boldsymbol{z}^{(1)} = \boldsymbol{z} \in \mathbb{R}^n$, Eq. (3.2) reduces to Eq. (1.81), thus Runge–Kutta is a special case of partitioned Runge–Kutta in which there is only one partition. Consequently, the result in the following section on the multisymplecticity of PRK methods also applies, in general, to RK methods.

For the multi-Hamiltonian PDE (1.6), discretisation in space with an r-stage PRK method gives the following set of equations:

$$\begin{aligned} \boldsymbol{Z}_{i,j}^{(\gamma)} &= \boldsymbol{z}_i^{(\gamma)} + \Delta x \sum_{k=1}^{r} a_{jk}^{(\gamma)} \partial_x \boldsymbol{Z}_{i,k}^{(\gamma)}, \quad j = 1, \ldots, r, \\ \boldsymbol{z}_{i+1}^{(\gamma)} &= \boldsymbol{z}_i^{(\gamma)} + \Delta x \sum_{j=1}^{r} b_j^{(\gamma)} \partial_x \boldsymbol{Z}_{i,j}^{(\gamma)}, \end{aligned} \tag{3.4}$$

for each γ, where the new variables $\partial_x \boldsymbol{Z}_{i,j}$ satisfy Eq. (1.6), i.e.,

$$\boldsymbol{K}\partial_t \boldsymbol{Z}_{i,j} + \boldsymbol{L}\partial_x \boldsymbol{Z}_{i,j} = \nabla_{\boldsymbol{z}} S(\boldsymbol{Z}_{i,j}). \tag{3.5}$$

Eqs. (3.4) and (3.5) form a differential-algebraic equation (DAE) for $\boldsymbol{Z}_{i,j}$ and $\boldsymbol{z}_i$. However, in this DAE there are no ODEs for the node values and the constraints only apply to $\boldsymbol{L}\boldsymbol{Z}_{i,j}$, not $\boldsymbol{Z}_{i,j}$. Furthermore, $\boldsymbol{L}$ may not have full rank, which may prevent one from obtaining a system of explicit ODEs for the $\boldsymbol{Z}_{i,j}$. This is explored in more detail in Section 3.3.

3.2 Multisymplecticity of PRK

Previous studies of the multi-Hamiltonian PDE (1.6) discretised in space and time with PRK methods have concluded that such discretisations satisfy a natural discrete approximation of the multisymplectic conservation law in Eq. (1.19) [38]. However, these studies use the same partitioning of the variables for both the spatial and time discretisations, which leads to a large number of cases to be considered, each with its own set of conditions to be satisfied. This choice of partitioning in each dimension is important as the conditions for the discretised equations to satisfy the discrete multisymplectic conservation law depend upon $\boldsymbol{K}$ and $\boldsymbol{L}$.

For example, given a multi-Hamiltonian PDE and a two-partition PRK discretisation in time with coefficients satisfying Eq. (1.90), if the PDE has no time derivatives of the variables in the second partition, then the discretisation is in fact a RK discretisation with the same coefficients as the first of the PRK pair, which will not, in general, satisfy Eq. (1.84).

To consider the most general case, we will now assume the finest possible partitioning of the variables, namely n partitions where, for each entry γ in $\boldsymbol{z}$ we have that $n_\gamma = 1$ and the partition $\boldsymbol{z}^{(\gamma)}$ consists simply of the variable z^γ.

Furthermore, we will consider the N-dimensional multi-Hamiltonian PDE (1.9). Because of this, we will require a slight extension of the notation in order to manage the indices in the following theorem and proof. To this end, for the following theorem, proof and corollary only, let $Z_{i_\alpha}^{\gamma,j_\delta}$ be the variable Z^γ located at the node i_α in the dimension x_α and at the internal stage j_δ in each other dimension x_δ.

The following theorem gives a much simpler set of conditions on the coefficients of the PRK methods such that the discretisation of a multi-Hamiltonian PDE by these methods results in a system of equations that satisfy a discrete multisymplectic conservation law. Since it immediately applies to any other partitioning of the variables by simply equating the $b_j^{(\gamma)}$ and $a_{ij}^{(\gamma)}$ coefficients of the appropriate partitions, this set of conditions encompasses all of the cases considered in previous studies.

Theorem 3.2.1. *A general multi-Hamiltonian PDE (1.9) discretised independently in each dimension, x_α, by an r_α-stage PRK method has a discrete multisymplectic conservation law that is the natural discrete approximation of Eq. (1.23) when, for each dimension x_α, the coefficients of the PRK method in that dimension satisfy the conditions*

$$\begin{aligned} b_j^{(\gamma)} &= b_j, \\ -a_{kj}^{(\gamma)} b_k^{(\beta)} - b_j^{(\gamma)} a_{jk}^{(\beta)} + b_j^{(\gamma)} b_k^{(\beta)} &= 0, \end{aligned} \tag{3.6}$$

for all j, k and pairs (β, γ) such that $\boldsymbol{K}^{(\alpha)}_{\beta\gamma} \neq 0$.

The discrete multisymplectic conservation law is then given by

$$\sum_{\alpha=1}^{N} \left(\sum_{\substack{\delta=1,\\ \delta\neq\alpha}}^{N} \sum_{j_\delta=1}^{r_\delta} \prod_{\substack{\epsilon=1,\\ \epsilon\neq\alpha}}^{N} \Delta x_\epsilon b^\epsilon_{j_\delta} \left(\omega^{\alpha,j_\delta}_{i_\alpha+1} - \omega^{\alpha,j_\delta}_{i_\alpha} \right) \right) = 0, \tag{3.7}$$

where $\omega^{\alpha,j_\delta}_{i_\alpha} = \frac{1}{2}\sum_{\beta,\gamma} \boldsymbol{K}^\alpha_{\beta\gamma} \mathrm{d}Z^{\gamma,j_\delta}_{i_\alpha} \wedge \mathrm{d}Z^{\beta,j_\delta}_{i_\alpha}$ and $b^\epsilon_{j_\delta}$ is the coefficient b_{j_δ} in the dimension x_ϵ.

Proof. For each dimension x_α, the difference in the 2-form $\omega^{\alpha,j_\delta}_{i_\alpha}$ between two adjacent grid points is given by

$$\begin{aligned}
\big(\omega^{\alpha,j_\delta}_{i_\alpha+1} - \omega^{\alpha,j_\delta}_{i_\alpha}\big) &= \frac{1}{2}\sum_{\beta,\gamma} \Big(\boldsymbol{K}^\alpha_{\beta\gamma} \mathrm{d}Z^{\gamma,j_\delta}_{i_\alpha+1} \wedge \mathrm{d}Z^{\beta,j_\delta}_{i_\alpha+1} - \boldsymbol{K}^\alpha_{\beta\gamma} \mathrm{d}Z^{\gamma,j_\delta}_{i_\alpha} \wedge \mathrm{d}Z^{\beta,j_\delta}_{i_\alpha} \Big) \\
&= \frac{1}{2}\sum_{\beta,\gamma} \boldsymbol{K}^\alpha_{\beta\gamma} \Big(\big(\mathrm{d}Z^{\gamma,j_\delta}_{i_\alpha} + \Delta x_\alpha \sum_j b^{(\gamma)}_j \partial_{x_\alpha} \mathrm{d}Z^{\gamma,j_\delta}_{i_\alpha,j}\big) \wedge \big(\mathrm{d}Z^{\beta,j_\delta}_{i_\alpha} + \Delta x_\alpha \sum_k b^{(\beta)}_k \partial_{x_\alpha} \mathrm{d}Z^{\beta,j_\delta}_{i_\alpha,k}\big) \\
&\quad - \mathrm{d}Z^{\gamma,j_\delta}_{i_\alpha} \wedge \mathrm{d}Z^{\beta,j_\delta}_{i_\alpha} \Big) \\
&= \frac{1}{2}\sum_{\beta,\gamma} \boldsymbol{K}^\alpha_{\beta\gamma} \Big(\Delta x_\alpha \big(\mathrm{d}Z^{\gamma,j_\delta}_{i_\alpha} \wedge \sum_k b^{(\beta)}_k \partial_{x_\alpha} \mathrm{d}Z^{\beta,j_\delta}_{i_\alpha,k} + \sum_j b^{(\gamma)}_j \partial_{x_\alpha} \mathrm{d}Z^{\gamma,j_\delta}_{i_\alpha,j} \wedge \mathrm{d}Z^{\beta,j_\delta}_{i_\alpha} \big) \\
&\quad + (\Delta x_\alpha)^2 \sum_{j,k} b^{(\gamma)}_j b^{(\beta)}_k \partial_{x_\alpha} \mathrm{d}Z^{\gamma,j_\delta}_{i_\alpha,j} \wedge \partial_{x_\alpha} \mathrm{d}Z^{\beta,j_\delta}_{i_\alpha,k} \Big) \\
&= \frac{1}{2}\sum_{\beta,\gamma} \boldsymbol{K}^\alpha_{\beta\gamma} \Big(\Delta x_\alpha \sum_k \big(\mathrm{d}Z^{\gamma,j_\delta}_{i_\alpha,k} - \Delta x_\alpha \sum_j a^{(\gamma)}_{kj} \partial_{x_\alpha} \mathrm{d}Z^{\gamma,j_\delta}_{i_\alpha,j} \big) \wedge b^{(\beta)}_k \partial_{x_\alpha} \mathrm{d}Z^{\beta,j_\delta}_{i_\alpha,k} \\
&\quad + \Delta x_\alpha \sum_j b^{(\gamma)}_j \partial_{x_\alpha} \mathrm{d}Z^{\gamma,j_\delta}_{i_\alpha,j} \wedge \big(\mathrm{d}Z^{\beta,j_\delta}_{i_\alpha,j} - \Delta x_\alpha \sum_k a^{(\beta)}_{jk} \partial_{x_\alpha} \mathrm{d}Z^{\beta,j_\delta}_{i_\alpha,k} \big) \\
&\quad + (\Delta x_\alpha)^2 \sum_{j,k} b^{(\gamma)}_j b^{(\beta)}_k \partial_{x_\alpha} \mathrm{d}Z^{\gamma,j_\delta}_{i_\alpha,j} \wedge \partial_{x_\alpha} \mathrm{d}Z^{\beta,j_\delta}_{i_\alpha,k} \Big) \\
&= \frac{1}{2}\sum_{\beta,\gamma} \boldsymbol{K}^\alpha_{\beta\gamma} \Big(\Delta x_\alpha \big(\sum_k b^{(\beta)}_k \mathrm{d}Z^{\gamma,j_\delta}_{i_\alpha,k} \wedge \partial_{x_\alpha} \mathrm{d}Z^{\beta,j_\delta}_{i_\alpha,k} + \sum_j b^{(\gamma)}_j \partial_{x_\alpha} \mathrm{d}Z^{\gamma,j_\delta}_{i_\alpha,j} \wedge \mathrm{d}Z^{\beta,j_\delta}_{i_\alpha,j} \big) \\
&\quad + (\Delta x_\alpha)^2 \sum_{j,k} \big(- a^{(\gamma)}_{kj} b^{(\beta)}_k - b^{(\gamma)}_j a^{(\beta)}_{jk} + b^{(\gamma)}_j b^{(\beta)}_k \big) \partial_{x_\alpha} \mathrm{d}Z^{\gamma,j_\delta}_{i_\alpha,j} \wedge \partial_{x_\alpha} \mathrm{d}Z^{\beta,j_\delta}_{i_\alpha,k} \Big)
\end{aligned} \tag{3.8}$$

$$= \Delta x_\alpha \sum_{\beta,\gamma,j} b_j^{(\gamma)} \boldsymbol{K}^\alpha_{\beta\gamma} \partial_{x_\alpha} \mathrm{d}Z^{\gamma,j_\delta}_{i_\alpha,j} \wedge \mathrm{d}Z^{\beta,j_\delta}_{i_\alpha,j}$$
$$+ \frac{1}{2}(\Delta x_\alpha)^2 \sum_{\beta,\gamma} \boldsymbol{K}^\alpha_{\beta\gamma} \sum_{j,k} \Big(- a^{(\gamma)}_{kj} b^{(\beta)}_k - b^{(\gamma)}_j a^{(\beta)}_{jk} + b^{(\gamma)}_j b^{(\beta)}_k \Big) \partial_{x_\alpha} \mathrm{d}Z^{\gamma,j_\delta}_{i_\alpha,j} \wedge \partial_{x_\alpha} \mathrm{d}Z^{\beta,j_\delta}_{i_\alpha,k}.$$

When $\boldsymbol{K}^\alpha_{\beta\gamma}$ is non-zero, the $(\Delta x_\alpha)^2$ term above is zero if

$$-a^{(\gamma)}_{kj} b^{(\beta)}_k - b^{(\gamma)}_j a^{(\beta)}_{jk} + b^{(\gamma)}_j b^{(\beta)}_k = 0, \qquad \text{for all } j,k. \tag{3.9}$$

Now, writing Eq. (1.10) in components and taking its wedge product with $\mathrm{d}z^\beta$ gives

$$\sum_{\alpha=1}^{N} \left(\sum_{\gamma} \boldsymbol{K}^\alpha_{\beta\gamma} \partial_{x_\alpha} \mathrm{d}z^\gamma \wedge \mathrm{d}z^\beta \right) = 0, \qquad \text{for all } \beta, \tag{3.10}$$

since $\boldsymbol{D}_{\boldsymbol{zz}} S(\boldsymbol{z})$ is symmetric. Thus, in general, multiplying the Δx_α term in the last line of Eq. (3.8) by $\Delta x_\epsilon b^\epsilon_{j_\delta}$ for all $\epsilon \neq \alpha$, then summing over $j_\delta \leq r_\delta$, $\delta \neq \alpha$ and α, gives

$$\sum_{\alpha=1}^{N} \left(\sum_{\substack{\delta=1,\\ \delta\neq\alpha}}^{N} \sum_{j_\delta=1}^{r_\delta} \prod_{\substack{\epsilon=1,\\ \epsilon\neq\alpha}}^{N} \Delta x_\epsilon b^\epsilon_{j_\delta} \left(\Delta x_\alpha \sum_{\beta,\gamma,j} b^{(\gamma)}_j \boldsymbol{K}^\alpha_{\beta\gamma} \partial_{x_\alpha} \mathrm{d}Z^{\gamma,j_\delta}_{i_\alpha,j} \wedge \mathrm{d}Z^{\beta,j_\delta}_{i_\alpha,j} \right) \right) = 0, \tag{3.11}$$

when $b^{(\gamma)}_j = b_j$ for all j and γ.

Therefore, if Eq. (3.6) holds then we can see from Eqs. (3.8) and (3.11) that the discrete multisymplectic conservation law (Eq. (3.7)) holds. □

For the multi-Hamiltonian PDE (1.6) in one time and one spatial dimension, Eq. (1.23) reduces to

$$\Delta x \sum_{j=1}^{r} b_j (\omega^{n+1}_{i,j} - \omega^n_{i,j}) + \Delta t \sum_{n=1}^{s} B_n (\kappa^{n,m}_{i+1} - \kappa^{n,m}_i) = 0, \tag{3.12}$$

where b_j are the coefficients of the PRK method in space and B_n are the coefficients of the PRK method in time.

The discrete multisymplectic conservation law (Eq. (3.12)) is an approximation to the integral

$$\int_{i\Delta x}^{(i+1)\Delta x} (\omega(x,(n+1)\Delta t) - \omega(x, n\Delta t))\,\mathrm{d}x + \int_{n\Delta t}^{(n+1)\Delta t} (\omega((i+1)\Delta x, t) - \omega(i\Delta x, t))\,\mathrm{d}t = 0, \tag{3.13}$$

which is the integral of Eq. (1.19) over the cell with lower left corner at $(i\Delta x, n\Delta t)$.

The multisymplecticity of a multi-Hamiltonian PDE discretised by a GRK method in each dimension is a special case of Theorem 3.2.1, where the coefficients $b_j^{(\gamma)}$ and $a_{ij}^{(\gamma)}$ in each dimension are equal for all the partitions $\boldsymbol{z}^{(\gamma)}$.

Often, a multi-Hamiltonian PDE is only discretised by PRK methods in a subset of the dimensions. This leads to the following corollary.

Corollary 3.2.2. *A general multi-Hamiltonian PDE (1.9) discretised independently in a subset, $\Sigma \subset \{1, \ldots, N\}$, of the dimensions of the PDE by PRK methods has a semi-discrete multisymplectic conservation law when, for each dimension x_α, $\alpha \in \Sigma$, the coefficients of the PRK method in that dimension satisfy Eq. (3.6) for all j, k and pairs (β, γ) such that $\boldsymbol{K}_{\beta\gamma}^{(\alpha)} \neq 0$. This semi-discrete multisymplectic conservation law is given by*

$$\sum_{\alpha=1}^{N} \left(\sum_{\substack{\delta \in \Sigma, \\ \delta \neq \alpha}} \sum_{j_\delta=1}^{r_\delta} \prod_{\substack{\epsilon \in \Sigma, \\ \epsilon \neq \alpha}} \Delta x_\epsilon b_{j_\delta}^{\epsilon} \left(\Omega^{\alpha, j_\delta} \right) \right) = 0, \tag{3.14}$$

where

$$\Omega^{\alpha, j_\delta} = \begin{cases} \omega_{i_\alpha+1}^{\alpha, j_\delta} - \omega_{i_\alpha}^{\alpha, j_\delta} & \text{if } \alpha \in \Sigma, \\ \partial_{x_\alpha} \omega^{\alpha, j_\delta} & \text{if } \alpha \neq \Sigma, \end{cases} \tag{3.15}$$

$\omega_{i_\alpha}^{\alpha, j_\delta} = \frac{1}{2} \sum_{\beta,\gamma} \boldsymbol{K}_{\beta\gamma}^{\alpha} \mathrm{d}Z_{i_\alpha}^{\gamma, j_\delta} \wedge \mathrm{d}Z_{i_\alpha}^{\beta, j_\delta}$, $\omega^{\alpha, j_\delta} = \frac{1}{2} \sum_{\beta,\gamma} \boldsymbol{K}_{\beta\gamma}^{\alpha} \mathrm{d}Z^{\gamma, j_\delta} \wedge \mathrm{d}Z^{\beta, j_\delta}$ *and* $b_{j_\delta}^{\epsilon}$ *is the coefficient* b_{j_δ} *in the dimension* x_ϵ.

The proof follows directly from the proof of Theorem 3.2.1 with the product and summation in the last step restricted to the dimensions that are discretised.

3.2.1 The importance of partitioning

For a given multi-Hamiltonian PDE, the choice of partitioning can be critical as to whether or not discretisation by a PRK method will give rise to a multisymplectic integrator.

For example, consider a multi-Hamiltonian PDE (1.6), discretised in space by a two-partition PRK method with coefficients satisfying Eq. (1.90). In general, the coefficients in each partition of the PRK method will not satisfy Eq. (1.84). Thus, in order to satisfy a discrete multisymplectic conservation law, the partitioning of the variables must be chosen such that the two-form κ contains no terms of the form $\mathrm{d}\boldsymbol{z}^{(1)} \wedge \mathrm{d}\boldsymbol{z}^{(1)}$ or $\mathrm{d}\boldsymbol{z}^{(2)} \wedge \mathrm{d}\boldsymbol{z}^{(2)}$.

For the canonical form (or Darboux normal form) of $\boldsymbol{L}$, given by

$$\boldsymbol{L} = \begin{bmatrix} \mathbf{0} & -\boldsymbol{I}_d & \mathbf{0} \\ \boldsymbol{I}_d & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \tag{3.16}$$

where $\boldsymbol{I}_d$ is the $d \times d$ identity matrix, the obvious choice of partitioning is to let $z^\gamma \in \boldsymbol{z}^{(1)}$ for $\gamma = 1, \ldots, d$, $z^\gamma \in \boldsymbol{z}^{(2)}$ for $\gamma = d+1, \ldots, 2d$ and $z^\gamma \in \boldsymbol{z}^{(3)}$ for $\gamma > 2d$. While there are 3 partitions in this partitioning, only 2 sets of RK coefficients are required for a PRK discretisation in space of the multi-Hamiltonian PDE as there are no spatial derivatives of the variables in $\boldsymbol{z}^{(3)}$ and the zero rows of $\boldsymbol{L}$ correspond to ODEs in time at each point in space.

In general, it is only possible to put one of $\boldsymbol{K}$ and $\boldsymbol{L}$ into its canonical form. So, if $\boldsymbol{K}$ is put into its canonical form, then the form of $\boldsymbol{L}$ will, in general, be more complicated and the choice of partitioning will also be more complicated. In this case, it may not be possible to partition the variables in space into less than n partitions in such a way that a PRK discretisation in space satisfies a multisymplectic conservation law.

Theorem 3.2.1 shows that if the partitioning of the variables in each dimension is chosen appropriately, then a PRK discretisation in each of these dimensions with coefficients satisfying Eq. (3.6) will result in a numerical integrator that formally satisfies a multisymplectic conservation law given by Eq. (3.7). However, this does not guarantee that the numerical integrator is well defined.

Such a discretisation may fail to form a well-defined multisymplectic integrator for the reasons labelled (i)–(iv) in Section 2.1. I have demonstrated, in Sections 2.2 and 2.3, how Gaussian RK discretisations of a multi-Hamiltonian PDE may suffer from these problems. However, in general, problems (iii) and (iv), which typically occur once boundary conditions are imposed, can be avoided by the use of discretisation methods that give rise to an explicit multisymplectic integrator.

In order to construct an explicit multisymplectic integrator, it is necessary (but not sufficient, as pointed out in Section 1.2.1) for the discretisation in each dimension to be both symplectic and explicit. For PDEs in one time and one spatial dimension, this condition means that a symplectic spatial discretisation must give rise to explicit ODEs in time (or vice-versa since time and space are treated on an equal footing).

This rules out discretisation by symplectic RK methods, such as GRK, since Eq. (1.84) implies that, for non-zero values of b_i, the values of a_{ii} are non-zero and thus such methods are necessarily implicit. However, no such restriction applies to PRK methods, and since the discretisation of a multi-Hamiltonian PDE by PRK methods with coefficients satisfying Eq. (3.6) gives a system of equations that satisfy a discrete multisymplectic conservation law, it is possible to construct an explicit (and hence, well-defined) multisymplectic integrator by discretising with such methods. For example, the well-known 5-point method obtained by applying leapfrog in space and time to the nonlinear wave equation, $u_{tt} - u_{xx} = -V'(u)$, gives the explicit multisymplectic integrator [13]:

$$\frac{1}{(\Delta t)^2}\left[\begin{array}{c} 1 \\ -2 \\ 1 \end{array}\right] u - \frac{1}{(\Delta x)^2}\left[\begin{array}{ccc} 1 & -2 & 1 \end{array}\right] u = -V'(u), \tag{3.17}$$

where we have used the notation of centred stencils. This approach will be explored in

Section 3.4 for the Lobatto IIIA–IIIB class of PRK methods, which I describe in the next section.

3.3 Lobatto IIIA–IIIB

The Lobatto IIIA–IIIB class of PRK methods are a collection of two-partition PRK discretisations whose coefficients are based upon Lobatto quadrature. For these methods, the coefficients $a_{ij}^{(1)}$, $a_{ij}^{(2)}$ and $b_j^{(1)} = b_j^{(2)} = b_j$ are determined by [16]:

$$\begin{aligned}
&B(\xi): && \sum_{i=1}^{r} b_i c_i^{k-1} = \frac{1}{k}, && \text{for } k \leq \xi, \\
&C(\xi): && \sum_{j=1}^{r} a_{ij}^{(1)} c_j^{k-1} = \frac{1}{k} c_i^k, && \text{for } i = 1, \dots, r \text{ and } k \leq \xi, \\
&D(\xi): && \sum_{i=1}^{r} b_i c_i^{k-1} a_{ij}^{(2)} = \frac{1}{k} b_j (1 - c_j^k), && \text{for } j = 1, \dots, r \text{ and } k \leq \xi,
\end{aligned} \tag{3.18}$$

for $\xi = r$, where the c_i are the zeros of the Lobatto quadrature polynomial,

$$\frac{\mathrm{d}^{r-2}}{\mathrm{d}x^{r-2}} \left(x^{r-1}(x-1)^{r-1} \right). \tag{3.19}$$

Lobatto IIIA is the class of RK methods whose coefficients, b_j and $a_{ij}^{(1)}$, are determined by Eq. (3.19), $B(\xi)$ and $C(\xi)$, whereas Lobatto IIIB is the class of RK methods whose coefficients, b_j and $a_{ij}^{(2)}$, are determined by Eq. (3.19), $B(\xi)$ and $D(\xi)$. When combined, they form the Lobatto IIIA–IIIB class of PRK methods, where Lobatto IIIA is applied to the variables in the first partition and Lobatto IIIB is applied to the variables in the second partition.

While the Lobatto IIIA and Lobatto IIIB classes of RK methods have each been known since the mid 1960s, their coefficients do not satisfy Eq. (1.84) and hence do not form symplectic or multisymplectic integrators when applied to Hamiltonian ODEs or multi-Hamiltonian PDEs. It was only discovered relatively recently that the Lobatto IIIA–IIIB class of PRK methods has coefficients that satisfy Eq. (1.90) [44, 70]. Thus, for a discretisation of Eq. (1.6) by Lobatto IIIA–IIIB methods, if the partitioning of the variables in each dimension can be chosen such that Eq. (3.6) is satisfied, then the resulting integrator will satisfy the discrete multisymplectic conservation law given by Eq. (3.12).

While the Lobatto IIIA–IIIB class of PRK methods have a lower order than PRK methods based on Gaussian or Radau quadrature, they have advantages that we will make use of in order to construct a well-defined multisymplectic integrator. In particular, the features of Lobatto IIIA–IIIB that will be seen to be important in the following section are that the coefficients are related in the following way:

$$a^{(1)}_{1j} = 0, \quad a^{(1)}_{rj} = b_j, \qquad \text{for all } j, \tag{3.20}$$

$$a^{(2)}_{ir} = 0, \quad a^{(2)}_{i1} = b_1, \qquad \text{for all } i, \tag{3.21}$$

and the $(r-2)\times(r-2)$ matrix $\boldsymbol{C}$ with entries

$$\boldsymbol{C}_{i-1,j-1} = \sum_{k,l} a^{(1)}_{ik}(b_l - \delta_{kl})a^{(2)}_{lj}, \qquad \text{for } 2 \le i,j \le r-1, \tag{3.22}$$

is invertible. The relations given in Eqs. (3.20) and (3.21) are a direct consequence of Eqs. (3.18) and (3.19) and give us three properties which will be required in our algorithm for constructing explicit ODEs in the next section. Firstly, from Eq. (3.20) we can see that for variables in the first partition, a node value is equal to the first stage value associated with that node and also equal to the last stage value associated with the previous node. Secondly, Eq. (3.21) gives us that both $\sum_j b_j a^{(2)}_{jr}$ and $b_1 - \sum_j b_j a^{(2)}_{j1}$ are zero. Lastly, Eqs. (3.20) and (3.21) together give

$$\sum_{k,l} a^{(1)}_{ik}(b_l - \delta_{kl})a^{(2)}_{lj} = 0 \qquad \text{if either } i \in \{1,r\} \text{ or } j \in \{1,r\}, \tag{3.23}$$

where δ_{kl} is the Kronecker delta. The invertibility of $\boldsymbol{C}$ can then be shown via the Frobenius inequality and will be used directly in the construction algorithm in the next section.

The coefficients for Lobatto IIIA–IIIB methods can be written succinctly as pairs of Butcher tableaux, which I give below for $r = 2$, 3 and 4.

$$r=2: \qquad \text{IIIA:} \quad \begin{array}{c|cc} 0 & 0 & 0 \\ 1 & \frac{1}{2} & \frac{1}{2} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}, \qquad \text{IIIB:} \quad \begin{array}{c|cc} 0 & \frac{1}{2} & 0 \\ 1 & \frac{1}{2} & 0 \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}. \tag{3.24}$$

Second order Lobatto IIIA–IIIB is often referred to as generalised leapfrog and, with the exception of the c_i values, is identical to Störmer–Verlet (Eq. (1.91)).

$$r=3: \qquad \text{IIIA:} \quad \begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{5}{24} & \frac{1}{3} & -\frac{1}{24} \\ 1 & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \\ \hline & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \end{array}, \qquad \text{IIIB:} \quad \begin{array}{c|ccc} 0 & \frac{1}{6} & -\frac{1}{6} & 0 \\ \frac{1}{2} & \frac{1}{6} & \frac{1}{3} & 0 \\ 1 & \frac{1}{6} & \frac{5}{6} & 0 \\ \hline & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \end{array}. \tag{3.25}$$

$$r=4: \qquad \text{IIIA:} \quad \begin{array}{c|cccc} 0 & 0 & 0 & 0 & 0 \\ \frac{5-\sqrt{5}}{10} & \frac{11+\sqrt{5}}{120} & \frac{25-\sqrt{5}}{120} & \frac{25-13\sqrt{5}}{120} & \frac{-1+\sqrt{5}}{120} \\ \frac{5+\sqrt{5}}{10} & \frac{11-\sqrt{5}}{120} & \frac{25+13\sqrt{5}}{120} & \frac{25+\sqrt{5}}{120} & \frac{-1-\sqrt{5}}{120} \\ 1 & \frac{1}{12} & \frac{5}{12} & \frac{5}{12} & \frac{1}{12} \\ \hline & \frac{1}{12} & \frac{5}{12} & \frac{5}{12} & \frac{1}{12} \end{array}, \tag{3.26}$$

$$r = 4: \qquad \text{IIIB:} \quad \begin{array}{c|cccc} 0 & \frac{1}{12} & \frac{-1-\sqrt{5}}{24} & \frac{-1+\sqrt{5}}{24} & 0 \\ \frac{5-\sqrt{5}}{10} & \frac{1}{12} & \frac{25+\sqrt{5}}{120} & \frac{25-13\sqrt{5}}{120} & 0 \\ \frac{5+\sqrt{5}}{10} & \frac{1}{12} & \frac{25+13\sqrt{5}}{120} & \frac{25-\sqrt{5}}{120} & 0 \\ 1 & \frac{1}{12} & \frac{11-\sqrt{5}}{24} & \frac{11+\sqrt{5}}{24} & 0 \\ \hline & \frac{1}{12} & \frac{5}{12} & \frac{5}{12} & \frac{1}{12} \end{array}. \tag{3.27}$$

For higher values of r, the tableaux become more complicated, but retain the properties given in Eqs. (3.20), (3.21) and (3.22). They can be easily computed from Eqs. (3.18) and (3.19).

The following derivation shows how Eq. (3.17) is obtained from a 2-stage Lobatto IIIA–IIIB discretisation in time and space of the nonlinear wave equation (with $\boldsymbol{z}$, $\boldsymbol{K}$, $\boldsymbol{L}$ and $S(\boldsymbol{z})$ given by Eq. (1.7)). For the spatial discretisation, the variables are partitioned into $\boldsymbol{z}^{(1)} = \{u, v\}$ and $\boldsymbol{z}^{(2)} = \{w\}$, which gives

$$\begin{aligned} W_{i,1} = W_{i,2} = W_{i,\frac{1}{2}} &= w_i + \frac{\Delta x}{2}(V'(U_{i,1} + \partial_t V_{i,1})) \\ &= w_{i+1} - \frac{\Delta x}{2}(V'(U_{i,2} + \partial_t V_{i,2})), \\ U_{i,1} &= u_i, \\ U_{i,2} &= u_{i+1} = u_i + \frac{\Delta x}{2}(W_{i,1} + W_{i,2}) = u_i + \Delta x W_{i,\frac{1}{2}}, \\ V_{i,1} &= v_i, \\ V_{i,2} &= v_{i+1}. \end{aligned} \tag{3.28}$$

Eliminating w_i, w_{i+1}, $W_{i,\frac{1}{2}}$, $U_{i,j}$ and $V_{i,j}$ gives

$$\frac{1}{(\Delta x)^2}(u_{i+2} - 2u_{i+1} + u_i) = V'(u_{i+1}) + \partial_t v_{i+1}. \tag{3.29}$$

Discretising this in time with the partitioning $\boldsymbol{z}^{(1)} = \{u, w\}$ and $\boldsymbol{z}^{(2)} = \{v\}$ gives

$$\begin{aligned} V_{i+1}^{n,1} = V_{i+1}^{n,2} = V_{i+1}^{n,\frac{1}{2}} &= v_{i+1}^n + \frac{\Delta t}{2}(\frac{1}{(\Delta x)^2}(U_{i+2}^{n,1} - 2U_{i+1}^{n,1} + U_i^{n,1})) - V'(U_{i+1}^{n,1}) \\ &= v_{i+1}^{n+1} - \frac{\Delta t}{2}(\frac{1}{(\Delta x)^2}(U_{i+2}^{n,2} - 2U_{i+1}^{n,2} + U_i^{n,2})) - V'(U_{i+1}^{n,2}), \\ U_{i+1}^{n,1} &= u_{i+1}^n, \\ U_{i+1}^{n,2} &= u_{i+1}^{n+1} = u_{i+1}^n + \frac{\Delta t}{2}(V_{i+1}^{n,1} + V_{i+1}^{n,2}) = u_{i+1}^n + \Delta t V_{i+1}^{n,\frac{1}{2}}. \end{aligned} \tag{3.30}$$

Eliminating v_{i+1}^n, v_{i+1}^{n+1}, $V_{i+1}^{n,\frac{1}{2}}$ and $U_{i+1}^{n,m}$ gives

$$\frac{1}{(\Delta t)^2}(u_{i+1}^{n+2} - 2u_{i+1}^{n+1} + u_{i+1}^n) - \frac{1}{(\Delta x)^2}(u_{i+2}^{n+1} - 2u_{i+1}^{n+1} + u_i^{n+1}) = -V'(u_{i+1}^{n+1}), \tag{3.31}$$

which is Eq. (3.17) written with indices.

3.4 Explicit Discretisation

We define an *explicit discretisation* in space as a discretisation for which the time derivatives of the dependent variables may be written explicitly in terms of the dependent variables. Their derivation may involve solving linear systems, but these must be independent of the PDE. An *explicit local discretisation* is an explicit discretisation for which these ODEs depend only on nearby values of the dependent variables.

In the one dimensional situation (i.e., time integration of a Hamiltonian ODE), the dependent variables in a PRK discretisation are the $\boldsymbol{z}_i$; Eq. (3.2) determines the value of the stage variables $\boldsymbol{Z}_{i,j}$ and defines a map from $\boldsymbol{z}_i$ to $\boldsymbol{z}_{i+1}$. In contrast, for situations where the dimension is greater than one (e.g., for PDEs of the form of Eq. (1.6)), if one applies a PRK discretisation in space, then the dependent variables will typically be the stage variables $\boldsymbol{Z}_{i,j}$, while the node variables $\boldsymbol{z}_i$ and the new variables $\partial_x \boldsymbol{Z}_{i,j}$ will be eliminated using the PDE to yield a set of ODEs in time for the $\boldsymbol{Z}_{i,j}$. As we shall see in the following theorem, this elimination depends upon the structure of not only $\boldsymbol{K}$ and $\boldsymbol{L}$, but also of $S(\boldsymbol{z})$.

Theorem 3.4.1. *Consider a multi-Hamiltonian PDE (1.6) where the $\boldsymbol{K}$ and $\boldsymbol{L}$ matrices have the following structure:*

$$\boldsymbol{K} = \begin{bmatrix} & -\boldsymbol{I}_{\frac{1}{2}(d_1+d_2)} & \\ \boldsymbol{I}_{\frac{1}{2}(d_1+d_2)} & & \\ & & \boldsymbol{0}_{d_1} \end{bmatrix}, \qquad \boldsymbol{L} = \begin{bmatrix} & & \boldsymbol{I}_{d_1} \\ & \boldsymbol{0}_{d_2} & \\ -\boldsymbol{I}_{d_1} & & \end{bmatrix} \tag{3.32}$$

where $d_1 = n - \operatorname{rank}(\boldsymbol{K})$, $d_2 = n - 2d_1 \le d_1$, $\boldsymbol{I}_d$ is the $d \times d$ identity matrix, $\boldsymbol{0}_d$ is the $d \times d$ zero matrix.

Let the variables $\boldsymbol{z}$ be partitioned into two partitions $\boldsymbol{z}^{(1)} \in \mathbb{R}^{d_1+d_2}$ and $\boldsymbol{z}^{(2)} \in \mathbb{R}^{d_1}$, where we denote the first d_1 components of $\boldsymbol{z}^{(1)}$ by $\boldsymbol{q}$, the last d_2 components of $\boldsymbol{z}^{(1)}$ by $\boldsymbol{v}$ and the components of $\boldsymbol{z}^{(2)}$ by $\boldsymbol{p}$ such that the PDE may be written as

$$\begin{bmatrix} & -\boldsymbol{I}_{\frac{1}{2}(d_1+d_2)} & \\ \boldsymbol{I}_{\frac{1}{2}(d_1+d_2)} & & \\ & & \boldsymbol{0}_{d_1} \end{bmatrix} \begin{bmatrix} \boldsymbol{q}_t \\ \boldsymbol{v}_t \\ \boldsymbol{p}_t \end{bmatrix} + \begin{bmatrix} & & \boldsymbol{I}_{d_1} \\ & \boldsymbol{0}_{d_2} & \\ -\boldsymbol{I}_{d_1} & & \end{bmatrix} \begin{bmatrix} \boldsymbol{q}_x \\ \boldsymbol{v}_x \\ \boldsymbol{p}_x \end{bmatrix} = \begin{bmatrix} \nabla_{\boldsymbol{q}} S(\boldsymbol{z}) \\ \nabla_{\boldsymbol{v}} S(\boldsymbol{z}) \\ \nabla_{\boldsymbol{p}} S(\boldsymbol{z}) \end{bmatrix}. \tag{3.33}$$

If the function $S(\boldsymbol{z})$ can be written in the form

$$S(\boldsymbol{z}) = T(\boldsymbol{p}) + V(\boldsymbol{q}) + \widehat{V}(\boldsymbol{v}), \tag{3.34}$$

where $T(\boldsymbol{p}) = \frac{1}{2}\boldsymbol{p}^t\beta\boldsymbol{p}$ and $\widehat{V}(\boldsymbol{v}) = \frac{1}{2}\boldsymbol{v}^T\alpha\boldsymbol{v}$ such that β and α are matrices with $|\beta| \neq 0$ and $|\alpha| \neq 0$, then applying an r-stage Lobatto IIIA–IIIB PRK discretisation in space to the PDE leads to a set of explicit local ODEs in time in the stage variables associated with $\boldsymbol{q}$.

Proof. Due to the form of $S(\boldsymbol{z})$, the central d_2 rows of Eq. (3.33) allow us to write entry i in $\boldsymbol{v}$ as

$$v_i = \sum_{j=1}^{d_2} (\alpha^{-1})_{i,j}\partial_t q_{j+\frac{1}{2}(d_1-d_2)}, \tag{3.35}$$

and hence

$$\partial_t v_i = \sum_{j=1}^{d_2} (\alpha^{-1})_{i,j}\partial_t^2 q_{j+\frac{1}{2}(d_1-d_2)}. \tag{3.36}$$

Substituting Eq. (3.36) into Eq. (3.33), we can eliminate the $\boldsymbol{v}$ variables in favour of higher order derivatives in time of the $\boldsymbol{q}$ variables. This lets us write Eq. (3.33) as

$$\boldsymbol{K}\boldsymbol{z}_t + \boldsymbol{L}\boldsymbol{z}_x - \mathcal{E}\boldsymbol{z}_{tt} = \nabla_{\boldsymbol{z}} S(\boldsymbol{z}), \tag{3.37}$$

where $\boldsymbol{z}$, $\boldsymbol{K}$, $\boldsymbol{L}$, $\mathcal{E}$ and $S(\boldsymbol{z})$ are the new vectors, matrices and functions given below:

$$\begin{gathered}
\boldsymbol{z} = \begin{bmatrix} \boldsymbol{q} \\ \boldsymbol{p} \end{bmatrix}, \quad \boldsymbol{K} = \begin{bmatrix} & & -\boldsymbol{I}_{\frac{1}{2}(d_1-d_2)} & \\ & \boldsymbol{0}_{d_2} & & \\ \boldsymbol{I}_{\frac{1}{2}(d_1-d_2)} & & & \\ & & & \boldsymbol{0}_{d_1} \end{bmatrix}, \\
\boldsymbol{L} = \begin{bmatrix} & \boldsymbol{I}_{d_1} \\ -\boldsymbol{I}_{d_1} & \end{bmatrix}, \quad \mathcal{E} = \begin{bmatrix} \boldsymbol{0}_{\frac{1}{2}(d_1-d_2)} & & & \\ & \alpha^{-1} & & \\ & & \boldsymbol{0}_{\frac{1}{2}(d_1-d_2)} & \\ & & & \boldsymbol{0}_{d_1} \end{bmatrix},
\end{gathered} \tag{3.38}$$

and $S(\boldsymbol{z}) = T(\boldsymbol{p}) + V(\boldsymbol{q})$.

Note that if $d_2 = 0$, then Eq. (3.33) and Eq. (3.37) are identical, i.e., $\widehat{V}(\boldsymbol{v}) \equiv 0$ and $\mathcal{E}$ is a $d_1 \times d_1$ matrix of zeros.

In order to complete the proof of this theorem, I shall now give a five step algorithm for constructing explicit local ODEs in time from an r-stage Lobatto IIIA–IIIB PRK discretisation of Eq. (3.37). However, before we begin, it is necessary to introduce the following notation, which extends the notation of Section 1.4.2:

- let z_i^η be the node variable in cell i for the entry η in $\boldsymbol{z}$,
- let $Z_{i,j}^\eta$ be the stage variable at stage j in cell i for the entry η in $\boldsymbol{z}$,
- let $\boldsymbol{Z}_i^\eta$ be the vector of stage variables in cell i for the entry η in $\boldsymbol{z}$,

- let $\boldsymbol{Z}_i$ be the tensor of stage variables for all values of η in cell i,
- let $\partial_t^n Z_{i,j}^\eta$ be the variable representing the first ($n = 1$) and second ($n = 2$) time derivatives of $Z_{i,j}^\eta$,
- let $\partial_{z^\eta} S(\boldsymbol{Z}_i)$ be the vector of stage values at cell i obtained by taking the derivative of the function $S(\boldsymbol{z})$ with respect to the entry η in $\boldsymbol{z}$,
- let $\boldsymbol{A}^{(1)}$ be the $r \times r$ matrix of a_{ij} values for Lobatto IIIA,
- let $\boldsymbol{A}^{(2)}$ be the $r \times r$ matrix of a_{ij} values for Lobatto IIIB,
- let $\boldsymbol{b}$ be the vector of length r of common b_j values for Lobatto IIIA and IIIB,
- let $\mathbf{1}$ be a vector of length r with all entries equal to 1.

Now, Eq. (3.37) discretised in space by an r-stage Lobatto IIIA–IIIB PRK discretisation results in the following system of implicit ODEs:

$$\boldsymbol{Q}_i^\eta = q_i^\eta \mathbf{1} + \Delta x \boldsymbol{A}^{(1)}(-\partial_{p^\eta} T(\boldsymbol{P}_i)), \tag{3.39a}$$

$$q_{i+1}^\eta = q_i^\eta + \Delta x \boldsymbol{b}^T(-\partial_{p^\eta} T(\boldsymbol{P}_i)), \tag{3.39b}$$

$$\boldsymbol{P}_i^\eta = p_i^\eta \mathbf{1} + \Delta x \boldsymbol{A}^{(2)}(\partial_{q^\eta} V(\boldsymbol{Q}_i) + g_i^\eta), \tag{3.39c}$$

$$p_{i+1}^\eta = p_i^\eta + \Delta x \boldsymbol{b}^T(\partial_{q^\eta} V(\boldsymbol{Q}_i) + g_i^\eta), \tag{3.39d}$$

for $1 \leq \eta \leq d_1$, where

$$g_i^\eta = \begin{cases} \partial_t \boldsymbol{Q}_i^{\eta + \frac{1}{2}(d_1+d_2)}, & 1 \leq \eta \leq \frac{1}{2}(d_1 - d_2), \\ -\partial_t \boldsymbol{Q}_i^{\eta - \frac{1}{2}(d_1+d_2)}, & \frac{1}{2}(d_1 + d_2) < \eta \leq d_1, \\ \sum_{\theta=1}^{d_2} (\alpha^{-1})_{\eta - \frac{1}{2}(d_1-d_2),\theta} \partial_t^2 \boldsymbol{Q}_i^{\theta + \frac{1}{2}(d_1-d_2)}, & \frac{1}{2}(d_1 - d_2) < \eta \leq \frac{1}{2}(d_1 + d_2). \end{cases} \tag{3.40}$$

It should be noted that for the simpler case where $d_2 = 0$, the third option for g_i^η vanishes.

Construction Algorithm:

Step 1:
Recall from Eq. (3.20) that, for the Lobatto IIIA discretisation, the first row of the coefficient matrix $\boldsymbol{A}^{(1)}$ is zero and the last row of $\boldsymbol{A}^{(1)}$ is $\boldsymbol{b}^T$.

Due to this property, we can see that the first row of Eq. (3.39a) gives $q_i^\eta = Q_{i,1}^\eta$ and comparing the last row of Eq. (3.39a) with Eq. (3.39b) gives $q_{i+1}^\eta = Q_{i,r}^\eta$. Furthermore, from these two identities we can conclude that $Q_{i,r}^\eta = Q_{i+1,1}^\eta$, $\partial_t Q_{i,r}^\eta = \partial_t Q_{i+1,1}^\eta$ and $\partial_t^2 Q_{i,r}^\eta = \partial_t^2 Q_{i+1,1}^\eta$.

Step 2:
Since $T(\boldsymbol{p}) = \frac{1}{2}\boldsymbol{p}^T \beta \boldsymbol{p}$ and $|\beta| \neq 0$ we have that $\boldsymbol{P}_i = \beta^{-1} \nabla_{\boldsymbol{p}} T(\boldsymbol{P}_i)$. Also, from $B(\xi)$ we

can see that all RK or PRK methods based on collocation have $\boldsymbol{b}^T\mathbf{1} = 1$. Therefore, we can substitute $\boldsymbol{P}_i^\eta$ from Eq. (3.39c) into Eq. (3.39b) and rearrange to get

$$p_i^\eta = -\frac{1}{\Delta x}\sum_{\zeta=1}^{d_1}\Big((\beta^{-1})_{\eta,\zeta}(Q_{i+1,1}^\zeta - Q_{i,1}^\zeta)\Big) - \Delta x \boldsymbol{b}^T\boldsymbol{A}^{(2)}(\partial_{q^\eta}V(\boldsymbol{Q}_i) + g_i^\eta). \tag{3.41}$$

Note that this rearrangement is possible since the matrix β operates on the index η, while $\boldsymbol{b}$ and $\boldsymbol{A}^{(2)}$ operate on the index j.

Step 3:
Substituting $\boldsymbol{P}_i^\eta$ from Eq. (3.39c) into Eq. (3.39a) and then substituting p_i^η from Eq. (3.41) into the resulting equation gives

$$\begin{aligned}
\boldsymbol{Q}_i^\eta &= Q_{i,1}^\eta\mathbf{1} - \Delta x\boldsymbol{A}^{(1)}\left(\sum_{\zeta=1}^{d_1}\beta_{\eta,\zeta}(\boldsymbol{P}_i^\zeta)\right)\\
&= Q_{i,1}^\eta\mathbf{1} - \Delta x\boldsymbol{A}^{(1)}\left(\sum_{\zeta=1}^{d_1}\beta_{\eta,\zeta}(p_i^\zeta\mathbf{1} + \Delta x\boldsymbol{A}^{(2)}(\partial_{q^\zeta}V(\boldsymbol{Q}_i) + g_i^\zeta))\right)\\
&= Q_{i,1}^\eta\mathbf{1} - \Delta x\boldsymbol{A}^{(1)}\Bigg(\sum_{\zeta=1}^{d_1}\beta_{\eta,\zeta}\Bigg(\Bigg[-\frac{1}{\Delta x}\sum_{\xi=1}^{d_1}\Big((\beta^{-1})_{\zeta,\xi}(Q_{i,r}^\xi - Q_{i,1}^\xi)\Big)\\
&\qquad - \Delta x\boldsymbol{b}^T\boldsymbol{A}^{(2)}(\partial_{q^\zeta}V(\boldsymbol{Q}_i) + g_i^\zeta)\Bigg]\mathbf{1} + \Delta x\boldsymbol{A}^{(2)}(\partial_{q^\zeta}V(\boldsymbol{Q}_i) + g_i^\zeta)\Bigg)\Bigg).
\end{aligned} \tag{3.42}$$

Rearranging and applying β^{-1} gives

$$\begin{aligned}
&\frac{1}{(\Delta x)^2}\sum_{\zeta=1}^{d_1}(\beta^{-1})_{\eta,\zeta}\left[\boldsymbol{Q}_i^\zeta - Q_{i,1}^\zeta\mathbf{1} - \boldsymbol{A}^{(1)}(Q_{i,r}^\zeta - Q_{i,1}^\zeta)\mathbf{1}\right]\\
&\qquad = \boldsymbol{A}^{(1)}\left[(\boldsymbol{b}^T\boldsymbol{A}^{(2)}(\partial_{q^\eta}V(\boldsymbol{Q}_i) + g_i^\eta))\mathbf{1} - \boldsymbol{A}^{(2)}(\partial_{q^\eta}V(\boldsymbol{Q}_i) + g_i^\eta)\right]\\
&\qquad = \boldsymbol{A}^{(1)}(\mathbf{1}\boldsymbol{b}^T - \boldsymbol{I})\boldsymbol{A}^{(2)}(\partial_{q^\eta}V(\boldsymbol{Q}_i) + g_i^\eta).
\end{aligned} \tag{3.43}$$

Now, the first and last rows of the left-hand side of Eq. (3.43) are zero, as are the first and last rows and columns of $\boldsymbol{A}^{(1)}(\mathbf{1}\boldsymbol{b}^T - \boldsymbol{I})\boldsymbol{A}^{(2)}$ due to Eqs. (3.20) and (3.21). Therefore, we denote rows 2 to $r-1$ of $\left[\boldsymbol{Q}_i^\zeta - Q_{i,1}^\zeta\mathbf{1} - \boldsymbol{A}^{(1)}(Q_{i,r}^\zeta - Q_{i,1}^\zeta)\mathbf{1}\right]$ by $\boldsymbol{d}_i^\zeta$, the block of $\boldsymbol{A}^{(1)}(\mathbf{1}\boldsymbol{b}^T - \boldsymbol{I})\boldsymbol{A}^{(2)}$ from $(2,2)$ to $(r-1,r-1)$ by $\boldsymbol{C}$ and rows 2 to $r-1$ of $\partial_{q^\eta}V(\boldsymbol{Q}_i) + g_i^\eta$ by $\boldsymbol{e}_i^\eta$. Then, noting from Eq. (3.22) that $\boldsymbol{C}$ has full rank, we can write

$$\frac{1}{(\Delta x)^2}\sum_{\zeta=1}^{d_1}(\beta^{-1})_{\eta,\zeta}\boldsymbol{C}^{-1}\boldsymbol{d}_i^\zeta = \boldsymbol{e}_i^\eta. \tag{3.44}$$

Recalling the definition of g_i^η, Eq. (3.44) immediately allows us to write down explicit formulas for $\partial_t Q_{i,k}^\eta$ in terms of $\boldsymbol{Q}_i$ for $1 < k < r$ and $1 \le \eta \le \frac{1}{2}(d_1 - d_2)$ or $\frac{1}{2}(d_1 + d_2) < \eta \le d_1$, and for $\partial_t^2 Q_{i,k}^\eta$ in terms of $\boldsymbol{Q}_i$ for $1 < k < r$ and $\frac{1}{2}(d_1 - d_2) < \eta \le \frac{1}{2}(d_1 + d_2)$.

Step 4:

Substituting p_i^η from Eq. (3.41) into Eq. (3.39d) for both p_i^η and p_{i+1}^η gives

$$-\frac{1}{(\Delta x)^2}\sum_{\zeta=1}^{d_1}(\beta^{-1})_{\eta,\zeta}(Q_{i+2,1}^\zeta - 2Q_{i+1,1}^\zeta + Q_{i,1}^\zeta) = \boldsymbol{b}^T\boldsymbol{A}^{(2)}(\partial_{q^\eta}V(\boldsymbol{Q}_{i+1}) + g_{i+1}^\eta) + (\boldsymbol{b}^T - \boldsymbol{b}^T\boldsymbol{A}^{(2)})(\partial_{q^\eta}V(\boldsymbol{Q}_i) + g_i^\eta) \quad (3.45)$$

for each η.

Of importance here, is that Eq. (3.45) does not involve the variables $\partial_t Q_{i+1,r}^\eta$ or $\partial_t^2 Q_{i+1,r}^\eta$ since the last entry of $\boldsymbol{b}^T\boldsymbol{A}^{(2)}$ is zero. Neither does it involve the variables $\partial_t Q_{i,1}^\eta$ or $\partial_t^2 Q_{i,1}^\eta$ since the first entry of $\boldsymbol{b}^T - \boldsymbol{b}^T\boldsymbol{A}^{(2)}$ is also zero. Both of these statements are a consequence of Eq. (3.21).

Step 5:

Substituting the formulas for $\partial_t Q_{i,k}^\eta$ and $\partial_t^2 Q_{i,k}^\eta$ found in Step 3 into Eq. (3.45) and recalling that $\partial_t Q_{i,r}^\eta = \partial_t Q_{i+1,1}^\eta$ and $\partial_t^2 Q_{i,r}^\eta = \partial_t^2 Q_{i+1,1}^\eta$, we can obtain explicit formulas for $\partial_t Q_{i+1,1}^\eta$ in terms of $\boldsymbol{Q}_i$ and $\boldsymbol{Q}_{i+1}$ for $1 \le \eta \le \frac{1}{2}(d_1 - d_2)$ and $\frac{1}{2}(d_1 + d_2) < \eta \le d_1$ and for $\partial_t^2 Q_{i+1,1}^\eta$ in terms of $\boldsymbol{Q}_i$ and $\boldsymbol{Q}_{i+1}$ for $\frac{1}{2}(d_1 - d_2) < \eta \le \frac{1}{2}(d_1 + d_2)$.

Thus, for each cell i in our grid, we have a system of explicit ODEs for either the first or second time derivatives of the stage variables $\boldsymbol{Q}_i$ in terms of local values of $\boldsymbol{Q}_i$. □

While the conditions on $\boldsymbol{K}$, $\boldsymbol{L}$ and $S(\boldsymbol{z})$ in the above theorem may at first appear restrictive, they allow several important equations such as the nonlinear wave and nonlinear Schrödinger equations. A notable exception is the Korteweg-de Vries equation for which $S(\boldsymbol{z})$ is not separable, i.e., it cannot be written in the form of Eq. (3.34). (The application of this theorem to a selection of PDEs, including the nonlinear wave, NLS and KdV equations, is explored in more detail in the next section.) It is also worth noting that the conditions on $\boldsymbol{K}$, $\boldsymbol{L}$ and $S(\boldsymbol{z})$ are the same as those required for the continuous system to be written as a system of PDEs in the variables $\boldsymbol{q}$ and are similar to those required for a separable Hamiltonian system to be written as a system of second order ODEs.

In the above theorem, $\boldsymbol{K}$ is canonical and, as mentioned in Section 3.2.1, writing any skew-symmetric $\boldsymbol{K}$ this way will, in general, complicate the structure of $\boldsymbol{L}$. However, if, after writing $\boldsymbol{K}$ this way, $\boldsymbol{L}$ has the following structure,

$$\boldsymbol{L} = \begin{bmatrix} & & \Lambda \\ & \boldsymbol{0}_{d_2} & \\ -\Lambda^T & & \end{bmatrix}, \quad (3.46)$$

for some $d_1 \times d_1$ matrix Λ with $|\Lambda| \neq 0$, then the following change of coordinates in the $\boldsymbol{p}$ variables can put $\boldsymbol{L}$ in the form given in Eq. (3.32).

Let

$$\hat{\boldsymbol{p}} = \Lambda \boldsymbol{p} \tag{3.47}$$

and

$$\widehat{T}(\hat{\boldsymbol{p}}) = T(\Lambda^{-1}\hat{\boldsymbol{p}}) = T(\boldsymbol{p}), \tag{3.48}$$

then

$$\nabla_{\hat{\boldsymbol{p}}} S(\boldsymbol{z}) = \nabla_{\hat{\boldsymbol{p}}} \widehat{T}(\hat{\boldsymbol{p}}) = \Lambda \nabla_{\boldsymbol{p}} T(\boldsymbol{p}) = \Lambda \beta \boldsymbol{p} = \Lambda \beta \Lambda^{-1} \hat{\boldsymbol{p}}. \tag{3.49}$$

The function $S(\boldsymbol{z})$ still has the desired structure,

$$S(\boldsymbol{z}) = V(\boldsymbol{q}) + \frac{1}{2}\hat{\boldsymbol{p}}^T(\Lambda\beta\Lambda^{-1})\hat{\boldsymbol{p}}. \tag{3.50}$$

The upper left $(d_1 + d_2) \times (d_1 + d_2)$ block of $\boldsymbol{L}$ being all zeros is fulfilled for PDEs that, when written as a first order system with a canonical $\boldsymbol{K}$, have no equations involving derivatives in both time and space of the same variable, i.e., $z_t^\eta + z_x^\eta = f(\boldsymbol{z})$ does not appear for any η. This is a feature that occurs frequently in "real world" multi-Hamiltonian PDEs.

3.4.1 Examples

Here I give several examples of common multi-Hamiltonian PDEs. For the PDEs that satisfy the requirements of Theorem 3.4.1, I give the ODEs that one obtains by applying the construction algorithm to those PDEs. For PDEs that do not satisfy the requirements of Theorem 3.4.1, I show why they fail and where the construction algorithm breaks down. I also give a PDE constructed so as to be as general as possible whilst still satisfying the conditions of Theorem 3.4.1.

Nonlinear wave equation

The first example is the nonlinear wave equation,

$$u_{tt} = u_{xx} - V'(u), \tag{3.51}$$

which can be written as a multi-Hamiltonian PDE in the form of Eq. (3.33) with $\boldsymbol{z}$, $\boldsymbol{K}$, $\boldsymbol{L}$ and $S(\boldsymbol{z})$ given by [8]

$$\boldsymbol{z} = \begin{bmatrix} u \\ v \\ w \end{bmatrix}, \quad \boldsymbol{K} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{L} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \tag{3.52}$$

and $S(\boldsymbol{z}) = V(u) + \frac{1}{2}v^2 - \frac{1}{2}w^2$.

Here, $d_1 = d_2 = 1$ with $\boldsymbol{z}^{(1)} = \{u, v\}$ and $\boldsymbol{z}^{(2)} = \{w\}$. We also have $\alpha = -\beta = 1$, thus we can see that $\boldsymbol{K}$, $\boldsymbol{L}$ and $S(\boldsymbol{z})$ satisfy the requirements of Theorem 3.4.1. Upon eliminating the variable v, we obtain the PDE (3.37) with

$$\boldsymbol{z} = \begin{bmatrix} u \\ w \end{bmatrix}, \quad \boldsymbol{K} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \boldsymbol{L} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \mathcal{E} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \tag{3.53}$$

and $S = V(u) - \frac{1}{2}w^2$.

Applying the construction algorithm for $r = 2$ gives the following pair of ODEs for each cell i,

$$\begin{aligned} \partial_t^2 U_{i,1} &= \frac{1}{(\Delta x)^2}(U_{i-1,1} - 2U_{i,1} + U_{i+1,1}) - V'(U_{i,1}), \\ \partial_t^2 U_{i,2} &= \partial_t^2 U_{i+1,1}. \end{aligned} \tag{3.54}$$

Recalling from Step 1 that $\boldsymbol{q}_i = \boldsymbol{Q}_{i,1}$ and noting that the last ODE is simply the first ODE of the next cell, it is convenient to drop the second ODE and rewrite the first ODE in terms of the node variable u_i, i.e.,

$$\partial_t^2 u_i = \frac{1}{(\Delta x)^2}(u_{i-1} - 2u_i + u_{i+1}) - V'(u_i). \tag{3.55}$$

Applying the construction algorithm for $r = 3$ gives the following triplet of ODEs for each cell i,

$$\begin{aligned} \partial_t^2 U_{i,1} &= \frac{1}{(\Delta x)^2}(-U_{i-1,1} + 8U_{i-1,2} - 14U_{i,1} + 8U_{i,2} - U_{i+1,1}) - V'(U_{i,1}), \\ \partial_t^2 U_{i,2} &= \frac{1}{(\Delta x)^2}(4U_{i,1} - 8U_{i,2} + 4U_{i+1,1}) - V'(U_{i,2}), \\ \partial_t^2 U_{i,3} &= \partial_t^2 U_{i+1,1}, \end{aligned} \tag{3.56}$$

which cannot be written in terms of the node variables alone.

NLS equation

The second example is the famous cubic-potential nonlinear Schrödinger (NLS) equation,

$$\mathrm{i}\psi_t + \psi_{xx} + 2|\psi|^2\psi = 0, \tag{3.57}$$

where $\psi \in \mathbb{C}$.

The NLS equation may be written as a multi-Hamiltonian PDE by first writing $\psi = p + \mathrm{i}q$ and separating the real and imaginary components of the equation to get

$$\begin{aligned} p_t &= -q_{xx} - 2(p^2 + q^2)q, \\ q_t &= p_{xx} + 2(p^2 + q^2)p. \end{aligned} \tag{3.58}$$

Then, these equations can be written in the form of Eq. (3.33) with $\boldsymbol{z}$, $\boldsymbol{K}$, $\boldsymbol{L}$ and $S(\boldsymbol{z})$ given by [42]

$$\boldsymbol{z} = \begin{bmatrix} p \\ q \\ v \\ w \end{bmatrix}, \quad \boldsymbol{K} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{L} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \tag{3.59}$$

and $S = -\frac{1}{2}(p^2 + q^2)^2 - \frac{1}{2}(v^2 + w^2)$.

Here we have $d_1 = 2$ and $d_2 = 0$ with the variables partitioned into $\boldsymbol{z}^{(1)} = \{p, q\}$ and $\boldsymbol{z}^{(2)} = \{v, w\}$. The function $S(\boldsymbol{z})$ can be written as Eq. (3.34) with $V(\boldsymbol{q}) = -\frac{1}{2}(p^2 + q^2)$ and $T(\boldsymbol{p}) = \frac{1}{2}\boldsymbol{p}^T \beta \boldsymbol{p}$ where

$$\beta = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \text{ and } \boldsymbol{p} = \begin{bmatrix} v \\ w \end{bmatrix}, \tag{3.60}$$

thus the NLS equation also satisfies the requirements of Theorem 3.4.1.

Applying the construction algorithm for an r-stage discretisation gives r ODEs for each element of $\boldsymbol{z}^{(1)}$ at cell i. As with the nonlinear wave equation, if we use the 2-stage discretisation then for each element of $\boldsymbol{z}^{(1)}$ at cell i we can drop the ODE for the second stage variable and write the ODE for the first stage variable in terms of the node variables. The resulting ODEs are

$$\begin{aligned} \partial_t p_i &= -\frac{1}{(\Delta x)^2}(q_{i-1} - 2q_i + q_{i+1}) - 2(p_i^2 + q_i^2)q_i, \\ \partial_t q_i &= \frac{1}{(\Delta x)^2}(p_{i-1} - 2p_i + p_{i+1}) + 2(p_i^2 + q_i^2)p_i. \end{aligned} \tag{3.61}$$

These are precisely the ODEs one obtains by applying second order finite differences in space to Eq. (3.58). The same statement applies for other PDEs that satisfy the conditions of Theorem 3.4.1, thus we note that 2-stage Lobatto IIIA–IIIB discretisation in space is equivalent to second order finite differences in space up to second order differences when applied to such a PDE. This similarity of the ODEs one obtains from applying a Lobatto IIIA–IIIB discretisation in space and the ODEs one obtains from applying finite differences in space to the multi-Hamiltonian PDE will be expanded upon in Section 3.4.2.

For $r = 3$ we obtain the following triplet of ODEs for each element of $\boldsymbol{z}^{(1)}$ at cell i,

$$\begin{aligned} \partial_t P_{i,1} &= -\frac{1}{(\Delta x)^2}(-Q_{i-1,1} + 8Q_{i-1,2} - 14Q_{i,1} + 8Q_{i,2} - Q_{i+1,1}) - 2(P_{i,1}^2 + Q_{i,1}^2)Q_{i,1}, \\ \partial_t P_{i,2} &= -\frac{1}{(\Delta x)^2}(4Q_{i,1} - 8Q_{i,2} + 4Q_{i+1,1}) - 2(P_{i,2}^2 + Q_{i,2}^2)Q_{i,2}, \\ \partial_t P_{i,3} &= \partial_t P_{i+1,1}, \end{aligned} \tag{3.62}$$

$$
\begin{aligned}
\partial_t Q_{i,1} &= \frac{1}{(\Delta x)^2}(-P_{i-1,1} + 8P_{i-1,2} - 14P_{i,1} + 8P_{i,2} - P_{i+1,1}) + 2(P_{i,1}^2 + Q_{i,1}^2)P_{i,1},\\
\partial_t Q_{i,2} &= \frac{1}{(\Delta x)^2}(4P_{i,1} - 8P_{i,2} + 4P_{i+1,1}) + 2(P_{i,2}^2 + Q_{i,2}^2)P_{i,2},\\
\partial_t Q_{i,3} &= \partial_t Q_{i+1,1}.
\end{aligned}
$$

Boussinesq equation

The third example is the "good" Boussinesq equation,

$$p_{tt} = (\varepsilon p_{xx} + V'(p))_{xx}, \tag{3.63}$$

which, when written as a multi-Hamiltonian PDE, shares the same $\boldsymbol{z}$, $\boldsymbol{z}^{(1)}$, $\boldsymbol{z}^{(2)}$, $\boldsymbol{K}$ and $\boldsymbol{L}$ as the NLS equation above [17]. The only difference is the function $S(\boldsymbol{z})$ which is given by $S(\boldsymbol{z}) = -V(p) + \frac{1}{2}(w^2 - \frac{1}{\varepsilon}v^2)$. (The class of Boussinesq equations includes a broad range of PDEs, some of which satisfy the conditions of Theorem 3.4.1.)

As before, the requirements of Theorem 3.4.1 are satisfied and applying the construction algorithm gives r ODEs for each element of $\boldsymbol{z}^{(1)}$ at cell i. For $r = 2$, we once again drop the ODEs for the second stage variables and write the first ODEs in terms of the node variables as

$$
\begin{aligned}
\partial_t p_i &= \frac{1}{(\Delta x)^2}(q_{i-1} - 2q_i + q_{i+1}),\\
\partial_t q_i &= \frac{\varepsilon}{(\Delta x)^2}(p_{i-1} - 2p_i + p_{i+1}) + V'(p).
\end{aligned} \tag{3.64}
$$

For $r = 3$ we get

$$
\begin{aligned}
\partial_t P_{i,1} &= \frac{1}{(\Delta x)^2}(-Q_{i-1,1} + 8Q_{i-1,2} - 14Q_{i,1} + 8Q_{i,2} - Q_{i+1,1}),\\
\partial_t P_{i,2} &= \frac{1}{(\Delta x)^2}(4Q_{i,1} - 8Q_{i,2} + 4Q_{i+1,1}),\\
\partial_t P_{i,3} &= \partial_t P_{i+1,1},\\
\partial_t Q_{i,1} &= \frac{\varepsilon}{(\Delta x)^2}(-P_{i-1,1} + 8P_{i-1,2} - 14P_{i,1} + 8P_{i,2} - P_{i+1,1}) + V'(P_{i,1}),\\
\partial_t Q_{i,2} &= \frac{\varepsilon}{(\Delta x)^2}(4P_{i,1} - 8P_{i,2} + 4P_{i+1,1}) + V'(P_{i,2}),\\
\partial_t Q_{i,3} &= \partial_t Q_{i+1,1}.
\end{aligned} \tag{3.65}
$$

Korteweg-de Vries (KdV) equation

The fourth example is the KdV equation,

$$u_t = V'(u)_x + \nu u_{xxx}, \tag{3.66}$$

which can be written in the form of Eq. (3.33) with [9]

$$\boldsymbol{z} = \begin{bmatrix} u \\ \phi \\ v \\ w \end{bmatrix}, \quad \boldsymbol{K} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{L} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \tag{3.67}$$

and with $S(\boldsymbol{z}) = -\frac{1}{2}uw - V(u) - \frac{1}{2\nu}v^2$. Here, $d_1 = 2$, $d_2 = 0$ and $\boldsymbol{z}$ is partitioned into $\boldsymbol{z}^{(1)} = \{u, \phi\}$ and $\boldsymbol{z}^{(2)} = \{v, w\}$.

While the $\boldsymbol{K}$ and $\boldsymbol{L}$ matrices have the required structure for Theorem 3.4.1, the function $S(\boldsymbol{z})$ does not. Specifically, the $-uw$ term in $S(\boldsymbol{z})$ prevents us from writing $T(\boldsymbol{p}) = \frac{1}{2}\boldsymbol{p}^T\beta\boldsymbol{p}$ and so Step 2 of the construction algorithm cannot be carried out.

For example, discretising the KdV equation with 2-stage Lobatto IIIA–IIIB gives

$$\begin{aligned} v_{i+\frac{1}{2}} &= v_{i-\frac{1}{2}} + \Delta x(\partial_t\phi_i - V'(u_i) - \frac{1}{4}(w_{i+\frac{1}{2}} + w_{i-\frac{1}{2}})), \\ w_{i+\frac{1}{2}} &= w_{i-\frac{1}{2}} - \Delta x\partial_t u_i, \\ -u_{i+1} &= -u_i - \Delta x\frac{1}{\nu}v_{i+\frac{1}{2}}, \\ -\phi_{i+1} &= -\phi_i - \Delta x\frac{1}{4}(u_i + u_{i+1}), \end{aligned} \tag{3.68}$$

where $u_i = U_{i,1}$, $u_{i+1} = U_{i,2}$, $\phi_i = \Phi_{i,1}$, $\phi_{i+1} = \Phi_{i,2}$, $v_{i+\frac{1}{2}} = V_{i,1} = V_{i,2}$ and $w_{i+\frac{1}{2}} = W_{i,1} = W_{i,2}$.

Introducing the operators D and M, where $Du_i = \frac{1}{\Delta x}(u_{i+1} - u_i)$ and $Mu_i = \frac{1}{2}(u_{i+1} + u_i)$, allows us to write this system as

$$\begin{aligned} Dv_{i-\frac{1}{2}} &= \partial_t\phi_i - V'(u_i) - \frac{1}{2}Mw_{i-\frac{1}{2}}, \\ Dw_{i-\frac{1}{2}} &= -\partial_t u_i, \\ -Du_i &= -\frac{1}{\nu}v_{i+\frac{1}{2}}, \\ -D\phi_i &= -\frac{1}{2}Mu_i. \end{aligned} \tag{3.69}$$

Then, eliminating all the variables other than the original variable u gives the implicit ODE

$$M\partial_t u_i = DV'(u_i) + \nu D^3 u_{i-1}. \tag{3.70}$$

In general, M is not invertible, thus further conditions are required (e.g., periodic boundary conditions with an odd number of grid points) to form a well-defined integrator from this implicit ODE.

This is none other than the *narrow box scheme*, introduced in [2] and derived as a finite volume scheme (and shown to be more accurate than the box scheme) in [3]. Thus, the narrow box scheme is shown to be multisymplectic.

Benjamin-Bona-Mahony (BBM) equation

The fifth example is the BBM equation [4],

$$u_t - \alpha u_{xxt} = V'(u)_x. \tag{3.71}$$

Writing this equation in the form of Eq. (1.6) with $\boldsymbol{z} = [u,\ \theta,\ \phi,\ w,\ \rho]^T$ gives

$$\boldsymbol{K} = \begin{bmatrix} 0 & \frac{\alpha}{2} & -\frac{1}{2} & 0 & 0 \\ -\frac{\alpha}{2} & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{L} = \begin{bmatrix} 0 & 0 & 0 & 0 & \frac{\alpha}{2} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ -\frac{\alpha}{2} & 0 & 0 & 0 & 0 \end{bmatrix} \tag{3.72}$$

and $S(\boldsymbol{z}) = uw - V(u) - \frac{\alpha}{2}\theta\rho$. The multi-Hamiltonian structure of the BBM equation presented here appears to be the first such occurrence in the literature.

Putting $\boldsymbol{K}$ into its Darboux normal form results in an $\boldsymbol{L}$ of the form

$$\boldsymbol{L} = \begin{bmatrix} \mathbf{0}_3 & \Lambda \\ -\Lambda^T & \mathbf{0}_2 \end{bmatrix}, \tag{3.73}$$

where Λ is a 3×2 matrix. The matrix $\boldsymbol{L}$ does not have the form of Eq. (3.46) and so it cannot be written in the form of Eq. (3.32) by applying a change of variables. Thus, the BBM equation does not satisfy the requirements of Theorem 3.4.1. However, partitioning $\boldsymbol{z}$ into $\boldsymbol{z}^{(1)} = \{u, \theta, \phi\}$ and $\boldsymbol{z}^{(2)} = \{w, \rho\}$, then discretising the BBM equation with 2-stage Lobatto IIIA–IIIB using the D and M notation gives

$$\begin{aligned} \frac{\alpha}{2} D\rho_{i-\frac{1}{2}} &= Mw_{i-\frac{1}{2}} - V'(u_i) - \frac{\alpha}{2}\partial_t\theta_i + \frac{1}{2}\partial_t\phi_i, \\ 0 &= -\frac{\alpha}{2} M\rho_{i-\frac{1}{2}} + \frac{\alpha}{2}\partial_t u_i, \\ -Dw_{i-\frac{1}{2}} &= -\frac{1}{2}\partial_t u_i, \\ D\phi_i &= Mu_i, \\ -\frac{\alpha}{2} Du_i &= -\frac{\alpha}{2} M\theta_i. \end{aligned} \tag{3.74}$$

Eliminating θ, ϕ, w and ρ gives the implicit ODE

$$(M^2 - \alpha D^2)\partial_t u_i = MDV'(u_i). \tag{3.75}$$

As with the KdV equation, the operator on the left-hand side cannot be locally inverted, although it is at least typically invertible.

Padé–II equation

The sixth example is the equation,

$$u_t - \alpha u_{xxt} = V'(u)_x + \nu u_{xxx}, \tag{3.76}$$

which contains a mixture of the third order derivatives found in the KdV and BBM equations. This equation is referred to as the Padé–II equation in [25] when $\nu = \frac{9}{10}$, $\alpha = \frac{19}{10}$ and $V(u) = -\frac{1}{2}u^2 - \frac{1}{6}u^3$. It can be written in the form of Eq. (1.6) with $\boldsymbol{z} = [u,\ \theta,\ \phi,\ w,\ \rho,\ v]^T$,

$$\boldsymbol{K} = \begin{bmatrix} 0 & \frac{\alpha}{2} & -\frac{1}{2} & 0 & 0 & 0 \\ -\frac{\alpha}{2} & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{L} = \begin{bmatrix} 0 & 0 & 0 & 0 & \frac{\alpha}{2} & \nu \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ -\frac{\alpha}{2} & 0 & 0 & 0 & 0 & 0 \\ -\nu & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{3.77}$$

and $S(\boldsymbol{z}) = uw - V(u) - \frac{\nu}{2}v^2 - \frac{\alpha}{2}\theta\rho$. This multi-Hamiltonian formulation of the Padé–II equation is, to the best of my knowledge, the first occurrence of such in the literature.

Putting $\boldsymbol{K}$ into its Darboux normal form results in an $\boldsymbol{L}$ of the form

$$\boldsymbol{L} = \begin{bmatrix} \mathbf{0}_3 & \Lambda \\ -\Lambda^T & \mathbf{0}_3 \end{bmatrix}, \tag{3.78}$$

where Λ is a 3×3 matrix with $\mathrm{rank}(\Lambda) = 2$. Thus, we cannot write $\boldsymbol{L}$ in the form of Eq. (3.32) and so the Padé–II equation does not satisfy the requirements of Theorem 3.4.1.

However, partitioning $\boldsymbol{z}$ into $\boldsymbol{z}^{(1)} = \{u, \theta, \phi\}$ and $\boldsymbol{z}^{(2)} = \{w, \rho, v\}$, then discretising the Padé–II equation with 2-stage Lobatto IIIA–IIIB using the D and M notation gives

$$\begin{aligned} \frac{\alpha}{2}D\rho_{i-\frac{1}{2}} + \nu D v_{i-\frac{1}{2}} &= M w_{i-\frac{1}{2}} - V'(u_i) - \frac{\alpha}{2}\partial_t\theta_i + \frac{1}{2}\partial_t\phi_i, \\ 0 &= -\frac{\alpha}{2}M\rho_{i-\frac{1}{2}} + \frac{\alpha}{2}\partial_t u_i, \\ -Dw_{i-\frac{1}{2}} &= -\frac{1}{2}\partial_t u_i, \\ D\phi_i &= M u_i, \\ -\frac{\alpha}{2}D u_i &= -\frac{\alpha}{2}M\theta_i, \\ -\nu D u_i &= -\nu v_{i-\frac{1}{2}}. \end{aligned} \tag{3.79}$$

Eliminating θ, ϕ, w, ρ and v gives the implicit ODE

$$(M^2 - \alpha D^2)\partial_t u_i = MDV'(u_i) + \nu M D^3 u_{i-1}. \tag{3.80}$$

As in the previous two examples, the operator on the left-hand side of Eq. (3.80) cannot be locally inverted.

A made-up example

The last example is contrived to satisfy the requirements of Theorem 3.4.1 and demonstrates the case when $d_2 \neq d_1$ and $d_2 \neq 0$. I have chosen $d_1 = 3$, $d_2 = 1$ and a multi-Hamiltonian PDE (1.6) with $\boldsymbol{z} = [q^1,\ q^2,\ q^3,\ v,\ p^1,\ p^2,\ p^3]^T$,

$$\boldsymbol{K} = \begin{bmatrix} 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{L} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{3.81}$$

and $S(\boldsymbol{z}) = V(\boldsymbol{q}) + \frac{1}{2}\boldsymbol{p}^T\beta\boldsymbol{p} + \frac{\alpha}{2}(v)^2$, where α is a constant and

$$\beta = \begin{bmatrix} 1 & 1 & -\frac{1}{2} \\ 1 & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{bmatrix}. \tag{3.82}$$

This corresponds to the PDE

$$\begin{aligned} \partial_t q^1 &= -2q^1_{xx} + 2q^2_{xx} + \partial_{q^3}V(\boldsymbol{q}), \\ \frac{1}{\alpha}\partial_t^2 q^2 &= -4q^1_{xx} + 3q^2_{xx} - 2q^3_{xx} - \partial_{q^2}V(\boldsymbol{q}), \\ \partial_t q^3 &= 4q^1_{xx} - 4q^2_{xx} + 2q^3_{xx} - \partial_{q^1}V(\boldsymbol{q}). \end{aligned} \tag{3.83}$$

Eliminating the variable v in favour of higher derivatives in time of q^2 gives Eq. (3.37) with

$$\boldsymbol{z} = \begin{bmatrix} q^1 \\ q^2 \\ q^3 \\ p^1 \\ p^2 \\ p^3 \end{bmatrix}, \quad \boldsymbol{K} = \begin{bmatrix} 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{L} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \end{bmatrix}, \tag{3.84}$$

$S(\boldsymbol{z}) = V(\boldsymbol{q}) + \frac{1}{2}\boldsymbol{p}^T\beta\boldsymbol{p}$, and the only non-zero entry of $\mathcal{E}$ given by $\mathcal{E}_{2,2} = \frac{1}{\alpha}$.

If we apply the construction algorithm for $r = 2$ then once again we can drop the

ODEs for the second stage variables and write the ODEs for the first stage variables in terms of the node variables giving the following ODEs at cell i,

$$\begin{aligned}
\partial_t q_i^1 &= \frac{1}{(\Delta x)^2}(-2q_{i-1}^1 + 2q_{i-1}^2 + 4q_i^1 - 4q_i^2 - 2q_{i+1}^1 + 2q_{i+1}^2) + \partial_{q^3} V(\boldsymbol{q}_i), \\
\partial_t^2 q_i^2 &= \frac{\alpha}{(\Delta x)^2}(-4q_{i-1}^1 + 3q_{i-1}^2 - 2q_{i-1}^3 + 8q_i^1 - 6q_i^2 + 4q_i^3 - 4q_{i+1}^1 + 3q_{i+1}^2 - 2q_{i+1}^3) \\
&\qquad - \alpha\partial_{q^2} V(\boldsymbol{q}_i), \\
\partial_t q_i^3 &= \frac{1}{(\Delta x)^2}(4q_{i-1}^1 - 4q_{i-1}^2 + 2q_{i-1}^3 - 8q_i^1 + 8q_i^2 - 4q_i^3 + 4q_{i+1}^1 - 4q_{i+1}^2 + 2q_{i+1}^3) \\
&\qquad - \partial_{q^1} V(\boldsymbol{q}_i).
\end{aligned} \tag{3.85}$$

3.4.2 A shortcut

An interesting feature of the system of explicit ODEs obtained in each of the examples above where Theorem 3.4.1 applies, is that the discretisation in space by Lobatto IIIA–IIIB only modifies the linear component of the multi-Hamiltonian PDE, that is, the discrete approximation of $\boldsymbol{L}\boldsymbol{z}_x$. The reason for this is that throughout the construction algorithm, the nonlinear components of the multi-Hamiltonian PDE always appear coupled to the time derivatives as the expression $\partial_{q^\eta} V(\boldsymbol{Q}_i) + g_i^\eta$.

Furthermore, in the examples above, the same pattern of coefficients arises from discretising different PDEs with the same order Lobatto IIIA–IIIB method. For example, with $r = 2$ the coefficients in the approximation of $\boldsymbol{q}_{xx}$ can be written as the centred stencil

$$\begin{bmatrix} 1 & -2 & 1 \end{bmatrix}, \tag{3.86}$$

while for $r = 3$ these coefficients can be written as the centred stencil

$$\begin{bmatrix} -1 & 8 & -14 & 8 & -1 \end{bmatrix} \tag{3.87}$$

for the first ODE and the stencil

$$\begin{bmatrix} 4 & -8 & 4 \end{bmatrix} \tag{3.88}$$

over a single cell for the second ODE.

This behaviour continues for higher values of r, e.g., for $r = 4$, the approximation of $\boldsymbol{q}_{xx}$ in the first ODE has coefficients that can be written as the centred stencil

$$\begin{bmatrix} 1 & \frac{1}{2}(25 - 15\sqrt(5)) & \frac{1}{2}(25 + 15\sqrt(5)) & -52 & \frac{1}{2}(25 + 15\sqrt(5)) & \frac{1}{2}(25 - 15\sqrt(5)) & 1 \end{bmatrix}, \tag{3.89}$$

while the second and third ODEs have coefficients that can be written as the stencils

$$\begin{bmatrix} 5 + 3\sqrt(5) & -20 & 10 & 5 - 3\sqrt(5) \end{bmatrix} \tag{3.90}$$

and

$$\begin{bmatrix} 5-3\sqrt(5) & 10 & -20 & 5+3\sqrt(5) \end{bmatrix} \tag{3.91}$$

over a single cell, respectively. As r increases, these patterns of the coefficients in the approximation of $\boldsymbol{q}_{xx}$ become increasingly complicated, yet for a given value of r, these patterns remain the same regardless of the PDE under consideration.

The reason these patterns of coefficients occur for different PDEs is due to Eqs. (3.44) and (3.45). For a given value of r, $\boldsymbol{C}$ and $\boldsymbol{d}_i^\zeta$ are fixed regardless of the PDE. Similarly, the coefficients $\boldsymbol{b}^T\boldsymbol{A}^{(2)}$ and $\boldsymbol{b}^T - \boldsymbol{b}^T\boldsymbol{A}^{(2)}$ in Eq. (3.45) are completely determined by r and Eqs. (3.18) and (3.19). Thus, when solving Eqs. (3.44) and (3.45) for g_i^η, the same weighting of the nearby stage variables occurs for $\boldsymbol{q}_{xx}$ for different PDEs.

For an r-stage discretisation, the approximation to q_{xx}^ζ at stage j (for $2 \le j \le r-1$) is given by

$$\begin{aligned} q_{xx}^\zeta((i+c_i)\Delta x) &\approx -\frac{1}{(\Delta x)^2}(\boldsymbol{C}^{-1}\boldsymbol{d}_i^\zeta)_{j-1} \\ &= \frac{1}{(\Delta x)^2}\sum_{k=2}^{r-1}(\boldsymbol{C}^{-1})_{j-1,k-1}((1-c_k)Q_{i,1}^\zeta - Q_{i,k}^\zeta + c_k Q_{i,r}^\zeta), \end{aligned} \tag{3.92}$$

where $\boldsymbol{C}$ is given by Eq. (3.22) and the c_k are the zeros of the Lobatto quadrature polynomial (Eq. (3.19)). For $j=1$, the approximation to q_{xx}^ζ is given by

$$\begin{aligned} q_{xx}^\zeta((i+1)\Delta x) \approx \frac{1}{2b_1(\Delta x)^2}\Bigg(\sum_{k=2}^{r-1}\Big((\boldsymbol{b}^T\boldsymbol{A}^{(2)})_k(\boldsymbol{C}^{-1}\boldsymbol{d}_{i+1}^\zeta)_{k-1} + (\boldsymbol{b}^T - \boldsymbol{b}^T\boldsymbol{A}^{(2)})_k(\boldsymbol{C}^{-1}\boldsymbol{d}_i^\zeta)_{k-1}\Big) \\ + (Q_{i+2,1}^\zeta - 2Q_{i+1,1}^\zeta + Q_{i,1}^\zeta)\Bigg), \end{aligned} \tag{3.93}$$

where b_1 is the first entry in $\boldsymbol{b}$.

This suggests the following shortcut:

1. Write the PDE with only terms of the form $\boldsymbol{z}_{xx}$ (i.e., no $\boldsymbol{z}_x$ terms).
2. Replace the $\boldsymbol{z}_{xx}$ terms with the PRK finite difference approximation of the desired order.

3.4.3 Boundary conditions

For many integrators, when various boundary conditions are applied to the system, the integrator either fails to remain well defined or it requires extra conditions to be so. This is referred to as problem (iii) in Section 2.1. Some examples of integrators that suffer from this problem are the implicit midpoint method [2] and higher order Gaussian Runge–Kutta methods as can be seen in Section 2.3. However, for integrators formed from the ODEs obtained by application of the construction algorithm in Section 3.3, the introduction of periodic, Dirichlet or Neumann boundary conditions presents no difficulties and the integrator remains well defined without any further restrictions, as is seen below.

Since the ODEs for the stage variables that are internal to each cell (i.e., $\partial_t \boldsymbol{Z}_{i,j}$ or $\partial_t^2 \boldsymbol{Z}_{i,j}$ for $2 \le j \le r-1$) do not couple together variables from outside the cell, each of the types of boundary conditions mentioned above has no impact on these ODEs. Thus, the only ODEs that need to be considered are those on the boundaries, i.e., on a spatial grid with N cells, the ODEs $\partial_t \boldsymbol{Z}_{0,1}$ or $\partial_t^2 \boldsymbol{Z}_{0,1}$ and $\partial_t \boldsymbol{Z}_{N,1}$ or $\partial_t^2 \boldsymbol{Z}_{N,1}$. For these ODEs, the various types of boundary conditions are handled as follows:

- For periodic boundary conditions, the spatial index is treated periodically with $\boldsymbol{Z}_{N,r}$ and its derivatives in time set equal to $\boldsymbol{Z}_{1,1}$ and its derivatives in time.

- For Dirichlet boundary conditions ($\boldsymbol{q}(t,x) = \textit{constant}$ on the boundaries), the values of $\boldsymbol{Q}_{1,1}$ and $\boldsymbol{Q}_{N,r}$ are simply given by the boundary conditions and hence, their time derivatives are zero. That is, $\partial_t \boldsymbol{Q}_{0,1} = \partial_t^2 \boldsymbol{Q}_{0,1} = \partial_t \boldsymbol{Q}_{N,1} = \partial_t^2 \boldsymbol{Q}_{N,1} = 0$.

- For Neumann boundary conditions ($\partial_x \boldsymbol{q}(t,x) = -\nabla_{\boldsymbol{p}} T(\boldsymbol{p}(t,x)) = -\beta \boldsymbol{p}(t,x) = \textit{constant}$ on the boundaries), the points outside the domain may be treated as phantom points.

  This can be seen by comparing Eq. (3.41) to Eq. (3.93). If $p_0 = 0$, then from Eq. (3.41) we have

$$\frac{1}{(\Delta x)^2}\sum_{\zeta=1}^{d_1}(\beta^{-1})_{\eta,\zeta}(Q_{1,1}^{\zeta} - Q_{0,1}^{\zeta}) = -\boldsymbol{b}^T \boldsymbol{A}^{(2)}(\partial_{q^\eta} V(\boldsymbol{Q}_i) + g_i^\eta)$$
$$= -\frac{1}{(\Delta x)^2}\sum_{k=2}^{r-1}(\boldsymbol{b}^T\boldsymbol{A}^{(2)})_k \sum_{\zeta=1}^{d_1}(\beta^{-1})_{\eta,\zeta}(\boldsymbol{C}^{-1}\boldsymbol{d}_0^{\zeta})_{k-1} - b_1(\partial_{q^\eta} V(\boldsymbol{Q}_{0,1}) + g_{0,1}^\eta), \quad (3.94)$$

  and so,

$$\partial_{q^\eta} V(\boldsymbol{Q}_{0,1}) + g_{0,1}^\eta =$$
$$-\sum_{\zeta=1}^{d_1}(\beta^{-1})_{\eta,\zeta}\frac{1}{b_1(\Delta x)^2}\left(Q_{1,1}^{\zeta} - Q_{0,1}^{\zeta} + \sum_{k=2}^{r-1}(\boldsymbol{b}^T\boldsymbol{A}^{(2)})_k(\boldsymbol{C}^{-1}\boldsymbol{d}_0^{\zeta})_{k-1}\right), \quad (3.95)$$

  that is, q_{xx}^{ζ} is approximated by

$$q_{xx}^{\zeta}(0) \approx \frac{1}{b_1(\Delta x)^2}\left(Q_{1,1}^{\zeta} - Q_{0,1}^{\zeta} + \sum_{k=2}^{r-1}(\boldsymbol{b}^T\boldsymbol{A}^{(2)})_k(\boldsymbol{C}^{-1}\boldsymbol{d}_0^{\zeta})_{k-1}\right). \quad (3.96)$$

  Now, recall that in Step 4 of the construction algorithm we made use of the last entry of $\boldsymbol{b}^T\boldsymbol{A}^{(2)}$ and the first entry of $\boldsymbol{b}^T - \boldsymbol{b}^T\boldsymbol{A}^{(2)}$ being zero. This can be further strengthened into the following lemma, which is proven in Appendix A.3.

Lemma 3.4.2. *The coefficients of the Lobatto IIIA–IIIB class of PRK methods satisfy*

$$\sum_{i=1}^{r} b_i A_{i,j}^{(2)} = b_{r+1-i} - \sum_{i=1}^{r} b_i A_{i,r+1-j}^{(2)}. \tag{3.97}$$

If we make the phantom points *ansatz* that $Q_{-1,1} = Q_{1,1}$ and $Q_{-1,j} = Q_{0,r+1-j}$ for $j = 2, \ldots, r-1$, then we can see from Lemma 3.4.2 that, for $i+1 = 0$, Eq. (3.93) reduces to Eq. (3.96).

For example, 3-stage Lobatto IIIA–IIIB applied to the NLS equation with Neumann boundary conditions, $\psi_x = 0$, applied to the left boundary as $v_1 = w_1 = 0$ leads to the following ODEs:

$$\begin{aligned} \partial_t P_{1,1} &= -\frac{1}{(\Delta x)^2}(-14Q_{1,1} + 16Q_{1,2} - 2Q_{2,1}) - 2(P_{1,1}^2 + Q_{1,1}^2)Q_{1,1}, \\ \partial_t Q_{1,1} &= \frac{1}{(\Delta x)^2}(-14P_{1,1} + 16P_{1,2} - 2P_{2,1}) + 2(P_{1,1}^2 + Q_{1,1}^2)P_{1,1}, \end{aligned} \tag{3.98}$$

which are equivalent to the first and fourth lines of Eq. (3.62), where the points outside the domain are treated as phantom points, i.e., $Q_{0,1} = Q_{2,1}$ and $Q_{0,2} = Q_{1,2}$.

It is worth mentioning that in this section I have only considered whether the system of explicit ODEs that one obtains from applying the construction algorithm in Section 3.3 to an appropriate PDE remains well defined when various types of boundary conditions are imposed. A full treatment of boundary conditions would require considerations such as how the stability of the method is affected, the existence of unwanted reflections from the boundaries, etc., which is beyond the scope of the thesis.

3.4.4 Other PRK discretisations

Finally, in this chapter, I would like to point out that although Theorem 3.4.1 is stated for the Lobatto IIIA–IIIB class of PRK discretisations, it applies equally well to any PRK discretisation satisfying Eqs. (3.20), (3.21) and (3.22). It remains an open question as to whether there are any other symplectic PRK discretisations with coefficients satisfying Eqs. (3.20), (3.21) and (3.22).

Chapter 4

Time Integration

In Section 3.2, it is shown that the discretisation of a multi-Hamiltonian PDE in space and time by PRK methods satisfying the conditions of Theorem 3.2.1 satisfies a discrete multisymplectic conservation law. In Section 3.4, it is shown that a Lobatto IIIA–IIIB discretisation in space of an appropriate multi-Hamiltonian PDE leads to explicit ODEs in time. Thus if an appropriate explicit PRK discretisation in time can be applied to these ODEs, one would obtain an explicit (and hence well-defined) multisymplectic integrator. An example of such an explicit multisymplectic integrator will be given in Section 4.3 for the NLS equation.

However, some multi-Hamiltonian PDEs lead to explicit ODEs that cannot be discretised in time by an appropriate PRK method to give an explicit multisymplectic integrator. Fortunately, there are other discretisation methods that may be applied in time to construct an explicit multisymplectic integrator from these ODEs. One such class of discretisation methods are known as symplectic splitting methods and will be discussed in Section 4.4 using the NLS equation as an example.

Much of the content of this chapter is contained in the recent paper "On multisymplecticity of partitioned Runge–Kutta and splitting methods" by myself, Robert McLachlan and Jason Frank [66], which was published in the Taylor and Francis journal *Int. J. Comput. Math.*.

4.1 Hamiltonian systems and explicit integration

A useful feature of the system of ODEs that one obtains from applying Theorem 3.4.1 to an appropriate multi-Hamiltonian PDE is that they may be written as a Hamiltonian system. E.g., for the Boussinesq equation and $r = 2$, the system of ODEs at node i can be written as

$$\partial_t \boldsymbol{z}_i = \boldsymbol{J}^{-1} \nabla_{\boldsymbol{z}_i} H_i \tag{4.1}$$

where

$$\boldsymbol{z}_i = \begin{bmatrix} q_i \\ p_i \end{bmatrix}, \quad \boldsymbol{J}^{-1} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tag{4.2}$$

and

$$H_i = \frac{1}{(\Delta x)^2}(-q_{i-1}q_i + q_i^2 - q_i q_{i+1} + \varepsilon p_{i-1}p_i - \varepsilon p_i^2 + \varepsilon p_i p_{i+1}) + V(p_i). \tag{4.3}$$

This can be done in general due to the form of the multi-Hamiltonian PDEs satisfying Theorem 3.4.1, i.e., Eq. (3.33). Writing such PDEs as

$$\boldsymbol{K}\boldsymbol{z}_t = \nabla_{\boldsymbol{z}} S(\boldsymbol{z}) - \boldsymbol{L}\boldsymbol{z}_x, \tag{4.4}$$

then discretising in space with Lobatto IIIA–IIIB and eliminating the $\boldsymbol{p}$ variables via Theorem 3.4.1 gives the Hamiltonian system of Eq. (4.1) for some function H_i.

In order to write such a system of ODEs as a Hamiltonian system, it is necessary that all of the ODEs that are second order in time be rewritten as pairs of first order ODEs. For example, in the made-up example in Section 3.4.1, the second line of Eq. (3.85) can be rewritten as

$$\begin{aligned} \partial_t q_i^2 &= \alpha v_i, \\ \partial_t v_i &= \frac{1}{(\Delta x)^2}(-4q_{i-1}^1 + 3q_{i-1}^2 - 2q_{i-1}^3 + 8q_i^1 - 6q_i^2 + 4q_i^3 - 4q_{i+1}^1 + 3q_{i+1}^2 - 2q_{i+1}^3) \\ &\quad - \alpha\partial_{q^2} V(\boldsymbol{q}_i), \end{aligned} \tag{4.5}$$

and the system can be written as Eq. (4.1) with $\boldsymbol{z}_i = [q_i^1,\ q_i^2,\ q_i^3,\ v_i]$ and

$$\begin{aligned} H_i = \frac{1}{(\Delta x)^2}\Big(&(-2q_{i-1}^1 + 2q_{i-1}^2 + 4q_i^1 - 4q_i^2 - 2q_{i+1}^1 + 2q_{i+1}^2)q_i^3 \\ &+ (-4q_{i-1}^1 + 4q_{i-1}^2 - 2q_{i-1}^3 + 4q_i^1 - 8q_i^2 - 4q_{i+1}^1 + 4q_{i+1}^2 - 2q_{i+1}^3)q_i^1 \\ &+ (4q_{i-1}^1 - 3q_{i-1}^2 + 2q_{i-1}^3 + 3q_i^2 + 4q_{i+1}^1 - 3q_{i+1}^2 + 2q_{i+1}^3)q_i^2\Big) \\ &+ \frac{\alpha}{2}(v_i)^2 + V(q_i^1, q_i^2, q_i^3). \end{aligned} \tag{4.6}$$

If the partitioning of the variables in the Hamiltonian, $(\boldsymbol{q}, \boldsymbol{p})$, is such that the Hamiltonian function is separable (i.e., it can be written as $H(\boldsymbol{z}) = T(\boldsymbol{p}) + V(\boldsymbol{q})$) then one can apply an explicit symplectic PRK discretisation (such as 2-stage Lobatto IIIA–IIIB) in time to obtain an explicit (and hence well-defined) multisymplectic integrator that is local and high order in space. Furthermore, explicit multisymplectic integrators that are also of high order in time may be obtained from such a Hamiltonian system by composition of 2-stage Lobatto IIIA–IIIB in time.

In the examples in Section 3.4.1, the systems of ODEs arising from the nonlinear wave equation and the Boussinesq equation both have separable Hamiltonians. Thus, explicit multisymplectic integrators that are local and of high order in space can be formed for these systems by applying an r-stage Lobatto IIIA–IIIB in space and a 2-stage Lobatto IIIA–IIIB in time.

If the Hamiltonian is not separable then other explicit time integrators may be applied, such as symplectic splitting methods [66], which may give superior performance (in terms of speed and stability) over implicit integrators. In Section 4.4, I will show how a symplectic splitting method in time may be applied to a multi-Hamiltonian PDE that has been appropriately discretised in space (such as with Lobatto IIIA–IIIB for PDEs satisfying Theorem 3.4.1) to construct an explicit multisymplectic integrator.

While the Hamiltonian of the explicit ODEs for NLS is non-separable, this non-separability only manifests as a scalar quadratic nonlinearity in the ODEs. For a 2-stage Lobatto IIIA–IIIB discretisation of these ODEs one obtains a pair of scalar quadratic equations which can be solved explicitly. Thus, for NLS, an explicit multisymplectic integrator can be formed by discretising these ODEs in time by 2-stage Lobatto IIIA–IIIB. The details of this will be given in Section 4.3.

Even if no explicit time integrator can be applied to the Hamiltonian system, there may be some benefits to having a spatial discretisation that gives rise to explicit ODEs, e.g., the ODEs may be less stiff than those obtained from an implicit discretisation.

4.2 Semi-discrete Multisymplectic Conservation Law for NLS

In the following sections I will describe several multisymplectic integrators based on the explicit ODEs obtained from a Lobatto IIIA–IIIB discretisation in space of an appropriate multi-Hamiltonian PDE. Throughout these sections I will be using the ODEs obtained from a 2-stage Lobatto IIIA–IIIB discretisation of the NLS equation as the primary example. Thus, let us first consider the semi-discrete multisymplectic conservation law that is satisfied by these ODEs.

The ODEs obtained from applying the construction algorithm in the proof of Theorem 3.4.1 to the NLS equation for a 2-stage Lobatto IIIA–IIIB discretisation in space are given in Eq. (3.61), which, for convenience, I restate here:

$$\begin{aligned} \partial_t p_i &= -\frac{1}{(\Delta x)^2}(q_{i-1} - 2q_i + q_{i+1}) - 2(p_i^2 + q_i^2)q_i, \\ \partial_t q_i &= \frac{1}{(\Delta x)^2}(p_{i-1} - 2p_i + p_{i+1}) + 2(p_i^2 + q_i^2)p_i. \end{aligned} \tag{4.7}$$

This can be written in the form of a Hamiltonian system (Eq. (4.1)) with $\boldsymbol{z} = [p_i,\ q_i]^T$ and

$$H_i = -\frac{1}{(\Delta x)^2}((q_{i-1} - q_i + q_{i+1})q_i + (p_{i-1} - p_i + p_{i+1})p_i) - \frac{1}{2}(p_i^2 + q_i^2)^2. \tag{4.8}$$

The semi-discrete multisymplectic conservation law for this discretisation of the NLS equation in space is found (in terms of the full set of variables in the multi-Hamiltonian PDE) from Eq. (3.14) with $\omega^1 = \mathrm{d}p \wedge \mathrm{d}q$, $\omega^2 = \mathrm{d}v \wedge \mathrm{d}p + \mathrm{d}w \wedge \mathrm{d}q$ and discretisation in the spatial dimension, x_2. That is,

$$\Delta x \sum_{j=1}^{2} \frac{1}{2} \partial_x \omega_{i,j} + \kappa_{i+1} - \kappa_i = 0, \tag{4.9}$$

where $\omega_{i,j} = \mathrm{d}P_{i,j} \wedge \mathrm{d}Q_{i,j}$ and $\kappa_i = \mathrm{d}v_i \wedge \mathrm{d}p_i + \mathrm{d}w_i \wedge \mathrm{d}q_i$. However, since the ODEs (4.7) are written in terms of a subset of the variables, it is desirable to write the semi-discrete multisymplectic conservation law in these variables too. This is accomplished by the following theorem:

Theorem 4.2.1. *The ODEs (4.7) satisfy the semi-discrete multisymplectic conservation law*

$$\partial_t(\mathrm{d}p_i \wedge \mathrm{d}q_i) + \frac{1}{(\Delta x)^2}\Big((\mathrm{d}p_{i+1} + \mathrm{d}p_{i-1}) \wedge \mathrm{d}p_i + (\mathrm{d}q_{i+1} + \mathrm{d}q_{i-1}) \wedge \mathrm{d}q_i\Big) = 0. \tag{4.10}$$

This theorem may be directly verified by substituting Eq. (4.7) into Eq. (4.10), however I give a proof in Appendix A.4 showing how Eq. (4.10) may be derived from the general form of the semi-discrete multisymplectic conservation law (Eq. (3.14)).

4.3 Integration by 2-stage Lobatto IIIA–IIIB

The first explicit multisymplectic integrator based on the ODEs (4.7) that I will consider is found by applying a 2-stage Lobatto IIIA–IIIB discretisation in time to these ODEs. Let the full set of variables in the multi-Hamiltonian form of the NLS equation be partitioned into $z^{(1)} = \{p, v\}$ and $z^{(2)} = \{q, w\}$ for the time discretisation. This partitioning of the variables satisfies the conditions of Theorem 3.2.1 and thus the resulting system of equations will formally satisfy a discrete multisymplectic conservation law.

In the ODEs (4.7) the variables v and w have been eliminated, so we are left with a standard p–q partitioning of the variables. With this partitioning of the variables, 2-stage Lobatto IIIA–IIIB is commonly known as generalized leapfrog, which, for an ODE $q_t = f(q,p)$, $p_t = g(q,p)$, is second order and takes the form [49]

$$\begin{aligned} q^{n+\frac{1}{2}} &= q^n + \frac{\Delta t}{2} f(q^{n+\frac{1}{2}}, p^n), \\ p^{n+1} &= p^n - \frac{\Delta t}{2}\Big(g(q^{n+\frac{1}{2}}, p^n) + g(q^{n+\frac{1}{2}}, p^{n+1})\Big), \\ q^{n+1} &= q^{n+\frac{1}{2}} + \frac{\Delta t}{2} f(q^{n+\frac{1}{2}}, p^{n+1}). \end{aligned} \tag{4.11}$$

In general it is an implicit method. Applying this to (4.7), one obtains the integrator that maps (p_i^n, q_i^n) to (p_i^{n+1}, q_i^{n+1}) in the following way:

$$
\begin{aligned}
q_i^{n+\frac{1}{2}} &= q_i^n + \frac{\Delta t}{2}\frac{1}{\Delta x^2}(p_{i-1}^n - 2p_i^n + p_{i+1}^n) + \Delta t((q_i^{n+\frac{1}{2}})^2 + (p_i^n)^2)p_i^n, \\
p_i^{n+1} &= p_i^n - \Delta t(q_{i-1}^{n+\frac{1}{2}} - 2q_i^{n+\frac{1}{2}} + q_{i+1}^{n+\frac{1}{2}}) - \Delta t((p_i^n)^2 + (p_i^{n+1})^2 + 2(q_i^{n+\frac{1}{2}})^2)q_i^{n+\frac{1}{2}}, \\
q_i^{n+1} &= q_i^{n+\frac{1}{2}} + \frac{\Delta t}{2}\frac{1}{\Delta x^2}(p_{i-1}^{n+1} - 2p_i^{n+1} + p_{i+1}^{n+1}) + \Delta t((q_i^{n+\frac{1}{2}})^2 + (p_i^{n+1})^2)p_i^{n+1}.
\end{aligned}
\tag{4.12}
$$

As mentioned above, the non-separable terms of the Hamiltonian enter into the integrator as scalar quadratic equations which can be solved explicitly. Furthermore, one of the solutions to the quadratic is $\mathcal{O}(1)$ while the other is of $\mathcal{O}(\Delta t^{-1})$, so the first solution is always taken.

As was noted in Section 3.4.2, applying a higher order Lobatto IIIA–IIIB discretisation in space to a multi-Hamiltonian PDE only modifies the approximation of the spatial derivative. Thus, with the above partitioning of the variables, a local, r-stage in space and second order in time, explicit multisymplectic integrator for NLS can be constructed by applying an r-stage Lobatto IIIA–IIIB discretisation in space and generalized leap-frog in time.

However, applying a higher order Lobatto IIIA–IIIB discretisation in time to Eq. (4.7) with this partitioning of the variables couples together the nonlinear terms and the resulting method is fully implicit. Instead, a higher order explicit integrator can be obtained by composition.

As with the semi-discrete multisymplectic conservation law in Theorem 4.2.1, it is desirable to write the fully discrete multisymplectic conservation law in terms of the variables in the integrator. The following theorem gives the discrete multisymplectic conservation law in the reduced set of variables.

Theorem 4.3.1. *For second order Lobatto IIIA–IIIB with partitioning $\{(p,q),(v,w)\}$ in space and $\{(p,v),(q,w)\}$ in time applied to the cubic NLS equation ($\mathrm{i}\psi_t + \psi_{xx} + 2|\psi|^2\psi = 0$) the discrete multisymplectic conservation law given in Eq. (3.7) can be written in terms of the local values of p and q as*

$$
\begin{aligned}
&\left(\frac{1}{\Delta t} + 2p_i^{n+1}q_i^{n+\frac{1}{2}}\right)\mathrm{d}p_i^{n+1} \wedge \mathrm{d}q_i^{n+\frac{1}{2}} - \left(\frac{1}{\Delta t} + 2p_i^n q_i^{n-\frac{1}{2}}\right)\mathrm{d}p_i^n \wedge \mathrm{d}q_i^{n-\frac{1}{2}} \\
&\quad + \frac{1}{\Delta x^2}\left((\mathrm{d}p_{i+1}^n + \mathrm{d}p_{i-1}^n) \wedge \mathrm{d}p_i^n + (\mathrm{d}q_{i+1}^{n+\frac{1}{2}} + \mathrm{d}q_{i-1}^{n+\frac{1}{2}}) \wedge \mathrm{d}q_i^{n+\frac{1}{2}}\right) = 0.
\end{aligned}
\tag{4.13}
$$

While, this theorem may also be directly verified by substituting Eq. (4.12) into Eq. (4.13), I give a proof in Appendix A.5 showing how Eq. (4.13) may be derived from the general form of the discrete multisymplectic conservation law (Eq. (3.7)).

4.4 Integration by symplectic splitting

Splitting methods for Hamiltonian ODEs were briefly introduced in Section 1.4.4. Here, I give an introduction to the use of splitting methods for multi-Hamiltonian PDEs in the construction of multisymplectic integrators.

Consider a multi-Hamiltonian PDE (1.6) where the vector field $\nabla_{\boldsymbol{z}} S(\boldsymbol{z})$ may be written more simply as a sum of vector fields, i.e.,

$$S(\boldsymbol{z}) = \sum_{j=1}^{N} S^{(j)}(\boldsymbol{z}). \tag{4.14}$$

Corresponding to this splitting of the vector field, introduce a splitting of the matrix $\boldsymbol{L}$,

$$\boldsymbol{L} = \sum_{j=1}^{N} \boldsymbol{L}^{(j)}, \tag{4.15}$$

such that each pair forms a multi-Hamiltonian PDE, i.e.,

$$\boldsymbol{K}\boldsymbol{z}_t + \boldsymbol{L}^{(j)}\boldsymbol{z}_x = \nabla_{\boldsymbol{z}} S^{(j)}(\boldsymbol{z}), \qquad \text{for } j = 1, \ldots, N. \tag{4.16}$$

If these subsystems are self-consistent, then the flow of each subsystem satisfies a different multisymplectic conservation law, i.e.,

$$\omega_t + \kappa_x^{(j)} = 0, \tag{4.17}$$

where $\kappa^{(j)} = \frac{1}{2}\boldsymbol{L}^{(j)}\mathrm{d}\boldsymbol{z}\wedge\mathrm{d}\boldsymbol{z}$. Since the density, ω, is the same in each conservation law, total symplecticity is conserved by the composition of the flows, which is the defining property of a symplectic splitting.

Suppose we can solve each of the subsystems over an interval Δt. Then the change in the 2-form, ω, for subsystem j is given by

$$\omega^{(j)} - \omega^{(j-1)} = -\int_0^{\Delta t} \kappa_x^{(j)}(\boldsymbol{z}^{(j)}(x,t))\,\mathrm{d}t, \tag{4.18}$$

where $\omega^{(0)} = \omega^0$ is the value of ω at the beginning of the interval. Combining the subsystems together gives

$$\omega^1 - \omega^0 = \omega^{(N)} - \omega^{(0)} = -\int_0^{\Delta t} \sum_{j=1}^{N} \kappa_x^{(j)}(\boldsymbol{z}^{(j)}(x,t))\,\mathrm{d}t, \tag{4.19}$$

which is of the same form as Eq. (1.19) but with the 2-forms in the integrals evaluated along different trajectories in phase space.

If the matrix $\boldsymbol{K}$ is degenerate, the projection of $\boldsymbol{L}\boldsymbol{z}_x = \nabla_{\boldsymbol{z}} S(\boldsymbol{z})$ upon the null space of

$\boldsymbol{K}$ will amount to constraints when the PDE is viewed as an infinite dimensional ODE. Often these constraints are removable by substitution as has been shown in Section 3.4. Nonetheless, it would appear necessary to preserve these constraints for each split subsystem. Furthermore, the part of $S(\boldsymbol{z})$ which defines the constraint should not be split apart from $\boldsymbol{L}\boldsymbol{z}_x$, since that can lead to strange relations like $\partial_x z^\gamma = 0$ for some components γ, that do not in general satisfy the initial conditions. It is necessary to check that the PDEs (4.16) are self-consistent for any proposed splitting.

In the following sections, two examples of multisymplectic integrators are given which are formed by applying splitting methods in time to the ODEs (4.7).

4.4.1 2-term (linear–nonlinear) splitting

The standard splitting of the cubic NLS equation is to separate the linear and nonlinear parts [36]. In terms of the original multi-Hamiltonian PDE, this splitting corresponds to taking $\boldsymbol{L}^{(1)} = \boldsymbol{L}$, $S^{(1)}(z) = -\frac{1}{2}(v^2 + w^2)$ for the first part, giving the linear PDE

$$\mathrm{i}\psi_t + \psi_{xx} = 0 \tag{4.20}$$

and taking $\boldsymbol{L}^{(2)} = \boldsymbol{0}$, $S^{(2)} = -\frac{1}{2}(p^2 + q^2)^2$ for the second part, giving the nonlinear PDE

$$\mathrm{i}\psi_t + 2|\psi|^2\psi = 0. \tag{4.21}$$

Note that since $\boldsymbol{L}^{(2)}$ is empty, the variables v and w drop out of this equation entirely and are undetermined; it is possible for the split subsystems to be singular in this way without destroying the method, as long as each subsystem is self-consistent.

After discretising in space, ψ_{xx} becomes $L\psi$ for some operator L. (In the original Fourier split-step method, L is the discrete Fourier derivative.) The time-Δt flow map of Eq. (4.20) is given by

$$\psi \mapsto \exp(\mathrm{i}\Delta t L)\psi, \tag{4.22}$$

while the time-Δt map of Eq. (4.21) is given by

$$\psi \mapsto \exp(\mathrm{i}\Delta t|\psi|^2)\psi. \tag{4.23}$$

For the 2-stage Lobatto IIIA–IIIB discretisation in space, this splitting corresponds to the splitting $H_i = H_{1,i} + H_{2,i}$ of the Hamiltonian Eq. (4.8), where

$$\begin{aligned} H_{1,i} &= -\frac{1}{(\Delta x)^2}((q_{i-1} - 2q_i + q_{i+1})q_i + (p_{i-1} - 2p_i + p_{i+1})p_i), \\ H_{2,i} &= -\frac{1}{2}(p_i^2 + q_i^2)^2. \end{aligned} \tag{4.24}$$

The operator L for the linear PDE (4.20) is represented by the centred stencil

$$-\frac{1}{(\Delta x)^2}\begin{bmatrix} 1 & -2 & 1 \end{bmatrix} \tag{4.25}$$

and thus the ODEs are given by

$$\begin{aligned} \partial_t p_i &= -\frac{1}{(\Delta x)^2}(q_{i-1} - 2q_i + q_{i+1}), \\ \partial_t q_i &= \frac{1}{(\Delta x)^2}(p_{i-1} - 2p_i + p_{i+1}). \end{aligned} \tag{4.26}$$

These ODEs are multisymplectic with conservation law Eq. (4.10); the only difference from Eq. (4.7) is that the nonlinear term, which only appears in $S(z)$, has been dropped. Therefore, the flow of the first part of the splitting satisfies Eq. (4.10) as well, although the differential forms are evaluated on solutions of Eq. (4.26) instead of (4.7).

One can derive a fully discrete multisymplectic conservation law for this part of the splitting, however there seems to be no unique or natural way to do this. For example, integrating Eq. (4.10) over one time step gives

$$\omega_i^1 - \omega_i^0 + \int_{t_0}^{t_0+\Delta t} \kappa_{i+1}(t) - \kappa_i(t)\,\mathrm{d}t = 0, \tag{4.27}$$

where $\omega_i = \mathrm{d}p_i \wedge \mathrm{d}q_i$ and $\kappa_i = \frac{1}{(\Delta x)^2}(\mathrm{d}p_i \wedge \mathrm{d}p_{i-1} + \mathrm{d}q_i \wedge \mathrm{d}q_{i-1})$. The explicit solution in terms of the initial conditions z^0 can now be substituted into $\kappa_i(t)$ to give a fully discrete multisymplectic conservation law. However, one could equally well express the solution in terms of the final state z^1, or some combination of z^0 and z^1, giving different discrete multisymplectic conservation laws.

Although (4.27) is local in space, the solution to Eq. (4.26) is not local in space, and therefore the resulting discrete multisymplectic conservation laws cannot be expected to be local, even though they are the exact solution of the local ODEs (4.10). (Note that for the PDE (4.20), the solution is local in space by an argument from the method of characteristics and thus, so is the multisymplectic conservation law; but after spatial discretisation the solution is no longer local in space and hence, neither is the discrete multisymplectic conservation law.)

The ODEs for the nonlinear PDE (4.21) are

$$\begin{aligned} \partial_t p_i &= -2(p_i^2 + q_i^2)q_i, \\ \partial_t q_i &= 2(p_i^2 + q_i^2)p_i. \end{aligned} \tag{4.28}$$

Since these ODEs are purely local in space and have a Hamiltonian given by the second line of Eq. (4.24), their flow satisfies $\partial_t \omega_i = 0$, and their time-Δt flow map satisfies $\omega_i^1 = \omega_i^0$.

The composition of the two parts of the splitting method therefore satisfies a non-local fully discrete multisymplectic conservation law. This conservation law is not a discrete analogue of the original semi-discrete multisymplectic conservation law (4.10), because during the second part, the 2-forms $\kappa_i(t)$ are changing in accordance with the solution of (4.28) instead of remaining fixed (as would be required to satisfy Eq. (4.10)). It will, however, be consistent with the original semi-discrete multisymplectic conservation law up to the order of the method.

As was observed in Section 3.4.2, applying a higher order Lobatto IIIA–IIIB discretisation in space only modifies the linear terms of the ODEs. The effect this has on the linear–nonlinear splitting is to modify the operator L such that more stage values are coupled together. While the details of Eqs. (4.26) and (4.28) and the semi-discrete multisymplectic conservation law change, their structure remains the same and the above splitting method applied to these ODEs will also satisfy a non-local discrete multisymplectic conservation law, however the details of evaluating the time-Δt flow map becomes more complicated.

4.4.2 3-term (real–imaginary–nonlinear) splitting

This splitting (which has been used for the linear Schrödinger equation [31]) differs from the one above in that the linear part of the above splitting is further split into real and imaginary parts. In this splitting the $\boldsymbol{L}^{(j)}$ matrices and corresponding $S^{(j)}(z)$ potentials are

$$\boldsymbol{L}^{(1)} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \qquad \boldsymbol{L}^{(2)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \qquad \boldsymbol{L}^{(3)} = \boldsymbol{0}, \tag{4.29}$$

$S^{(1)}(z) = -\frac{1}{2}v^2$, $S^{(2)}(z) = -\frac{1}{2}w^2$ and $S^{(3)}(z) = -\frac{1}{2}(p^2+q^2)^2$.

For the 2-stage Lobatto IIIA–IIIB discretisation in space, this splitting corresponds to the splitting $H_i = H_{1,i} + H_{2,i} + H_{3,i}$ of the Hamiltonian Eq. (4.8), where

$$\begin{aligned} H_{1,i} &= -\frac{1}{(\Delta x)^2}(p_{i-1} - 2p_i + p_{i+1})p_i, \\ H_{2,i} &= -\frac{1}{(\Delta x)^2}(q_{i-1} - 2q_i + q_{i+1})q_i, \\ H_{3,i} &= -\frac{1}{2}(p_i^2 + q_i^2)^2. \end{aligned} \tag{4.30}$$

In the solution to the first part, the value of p_i remains constant, as does the value of q_i in the solution to the second part. Therefore, the above splitting yields the following three time-Δt flow maps.

- Flow map 1:

$$
\begin{aligned}
p_i^{(1)} &= p_i^{(0)}, \\
q_i^{(1)} &= q_i^{(0)} + \frac{\Delta t}{(\Delta x)^2}(p_{i-1}^{(1)} - 2p_i^{(1)} + p_{i+1}^{(1)}),
\end{aligned}
\tag{4.31}
$$

for which the exact time-Δt flow map satisfies

$$
\mathrm{d}p_i^{(1)} \wedge \mathrm{d}q_i^{(1)} = \mathrm{d}p_i^{(0)} \wedge \mathrm{d}q_i^{(0)} + \frac{\Delta t}{(\Delta x)^2}(\mathrm{d}p_i^{(1)} \wedge (\mathrm{d}p_{i-1}^{(1)} + \mathrm{d}p_{i+1}^{(1)})).
\tag{4.32}
$$

- Flow map 2:

$$
\begin{aligned}
p_i^{(2)} &= p_i^{(1)} - \frac{\Delta t}{(\Delta x)^2}(q_{i-1}^{(1)} - 2q_i^{(1)} + q_{i+1}^{(1)}), \\
q_i^{(2)} &= q_i^{(1)},
\end{aligned}
\tag{4.33}
$$

for which the exact time-Δt flow map satisfies

$$
\mathrm{d}p_i^{(2)} \wedge \mathrm{d}q_i^{(2)} = \mathrm{d}p_i^{(1)} \wedge \mathrm{d}q_i^{(1)} - \frac{\Delta t}{(\Delta x)^2}((\mathrm{d}q_{i-1}^{(1)} + \mathrm{d}q_{i+1}^{(1)}) \wedge \mathrm{d}q_i^{(1)}).
\tag{4.34}
$$

- Flow map 3:

$$
\begin{aligned}
\begin{pmatrix} q_i^{(3)} \\ p_i^{(3)} \end{pmatrix} &= \begin{bmatrix} \cos \alpha_i \Delta t & \sin \alpha_i \Delta t \\ -\sin \alpha_i \Delta t & \cos \alpha_i \Delta t \end{bmatrix} \begin{pmatrix} q_i^{(2)} \\ p_i^{(2)} \end{pmatrix}, \\
\alpha_i &:= 2((q_i^{(2)})^2 + (p_i^{(2)})^2),
\end{aligned}
\tag{4.35}
$$

for which the exact time-Δt flow map satisfies

$$
\mathrm{d}p_i^{(3)} \wedge \mathrm{d}q_i^{(3)} = \mathrm{d}p_i^{(2)} \wedge \mathrm{d}q_i^{(2)}.
\tag{4.36}
$$

Substituting (4.32) into (4.34) and the result into (4.36), we can see that the composition of the three parts exactly satisfies

$$
\frac{1}{\Delta t}(\mathrm{d}p_i^{(3)} \wedge \mathrm{d}q_i^{(3)} - \mathrm{d}p_i^{(0)} \wedge \mathrm{d}q_i^{(0)}) + \frac{1}{(\Delta x)^2}((\mathrm{d}p_{i-1}^{(1)} + \mathrm{d}p_{i+1}^{(1)}) \wedge \mathrm{d}p_i^{(1)} + (\mathrm{d}q_{i-1}^{(1)} + \mathrm{d}q_{i+1}^{(1)}) \wedge \mathrm{d}q_i^{(1)}) = 0,
\tag{4.37}
$$

which is a fully discrete, local version of the semi-discrete multi-symplectic conservation law, Eq. (4.10). Unlike the solution to Eq. (4.26), the solutions to Eqs. (4.31) and (4.33) are both local since in Eq. (4.31) the value of p_i remains constant and in Eq. (4.33) the value of q_i remains constant. Furthermore, since the solution to Eq. (4.35) is also local and all three parts of the method are explicit, the overall method is explicit and local.

By the BCH theorem, this method is first order in time. As with the 2-term splitting in the previous section, this 3-term splitting may be applied to the Hamiltonian system obtained by applying a higher order Lobatto IIIA–IIIB discretisation in space. Doing so only modifies the update to the q_i component of the first flow map and the update to the p_i component of the second flow map, thus the method remains explicit and local, however it will satisfy a different discrete multisymplectic conservation law. Similarly, a higher-order method in time is attainable by using a more complicated composition of these three flow maps, but again, the resulting method will satisfy a different discrete multisymplectic conservation law.

The two multisymplectic integrators in this section and the previous one demonstrate how a well-defined explicit multisymplectic integrator may be constructed from the ODEs that one obtains by applying a Lobatto IIIA–IIIB discretisation in space and a splitting method in time to the NLS equation. For other multi-Hamiltonian PDEs that satisfy the conditions of Theorem 3.4.1, well-defined explicit multisymplectic integrators may be formed by an appropriate splitting of the Hamiltonian ODEs. These multisymplectic integrators will be local if each part of the splitting method can be solved locally and the discrete multisymplectic conservation laws that they satisfy may or may not be a discrete analogue of the semi-discrete multisymplectic conservation law that the Hamiltonian ODEs satisfy, however, they will, at least, agree with the semi-discrete multisymplectic conservation law up to the order of the splitting method.

4.5 Conservation laws

NLS has three basic conservation laws that can be derived in a unified way from a multisymplectic form of Noether's theorem: energy from the symmetry $t \mapsto t + c$, momentum from the symmetry $x \mapsto x + c$, and norm from the phase symmetry $\psi \mapsto e^{i\theta}\psi$. Discrete versions of these can be preserved if the symmetry is preserved. Spatial discretisation by a fixed grid destroys the spatial translation symmetry but not the time or phase symmetries, hence the semi-discrete system (Eq. (4.7)) has semi-discrete conservation laws for energy and norm but not for momentum. Time discretisation destroys the time symmetry, thus no method will have a fully discrete energy conservation law.

That leaves the norm conservation law, which for the PDE (3.58) is

$$(q^2 + p^2)_t + (pq_x - qp_x)_x = 0, \tag{4.38}$$

with associated conserved quantity (subject to suitable boundary conditions) $\int (q^2+p^2)\,\mathrm{d}x$ $(= \|\psi\|_2^2)$. The semi-discrete conservation law associated with Eq. (4.7) is

$$(q_i^2 + p_i^2)_t + \frac{1}{(\Delta x)^2}(p_i q_{i+1} - q_i p_{i+1} - p_{i-1} q_i + q_{i-1} p_i) = 0 \tag{4.39}$$

with conserved quantity $\sum_i (q_i^2 + p_i^2)$.

Time discretisation by the Lobatto IIIA–IIIB method (Section 4.3) and the real–imaginary splitting (Section 4.4.2) do not preserve the phase symmetry because of the splitting across the p–q variables. Those methods therefore do not have a discrete norm conservation law and do not conserve $\sum_i (q_i^2 + p_i^2)$. Time discretisation by linear–nonlinear splitting (Section 4.4.1) does preserve the phase symmetry. The first (linear) step obeys Eq. (4.39) exactly, while the second (nonlinear) step obeys $(q_i^2 + p_i^2)_t = 0$. Therefore this method has a discrete norm conservation law in the same sense in which it has a discrete multisymplectic conservation law: integrating Eq. (4.39) and substituting the solution gives a non-local conservation law. Such a law is, however, sufficient to conserve $\sum_i (q_i^2 + p_i^2)$.

Finally, time discretisation of Eq. (4.7) by the well-known implicit midpoint method also preserves the phase symmetry. This method is well known to preserve the total norm; here, we show that it also satisfies a discrete conservation law:

$$\begin{aligned}((q_i^{n+1})^2 + (p_i^{n+1})^2) - ((q_i^n)^2 + (p_i^n)^2) \\ &= (q_i^{n+1} - q_i^n)(q_i^{n+1} + q_i^n) + (p_i^{n+1} - p_i^n)(p_i^{n+1} + p_i^n) \\ &= \Delta t(2\bar{q}_i(D^2\bar{p}_i + 2(\bar{q}_i^2 + \bar{p}_i^2)\bar{p}_i) + 2\bar{p}_i(-D^2\bar{q}_i - 2(\bar{q}_i^2 + \bar{p}_i^2)\bar{q}_i)) \\ &= 2\Delta t(\bar{q}_i\bar{p}_{i-1} - \bar{q}_{i+1}\bar{p}_i + \bar{q}_i\bar{p}_{i+1} - \bar{q}_{i-1}\bar{p}_i),\end{aligned} \tag{4.40}$$

where $\bar{q}_i = \frac{1}{2}(q_i^{n+1} + q_i^n)$, $\bar{p}_i = \frac{1}{2}(p_i^{n+1} + p_i^n)$ and D^2 is the central difference approximation of ∂_{xx}. Hence, the midpoint rule conserves $\sum_i (q_i^2 + p_i^2)$.

It is striking that the non-local and implicit methods have a discrete norm conservation law, while the local and explicit methods do not.

Chapter 5

Dispersion and Order

Many of the important properties of a numerical integrator (such as the stability and the sign of the group velocity) may be determined by considering the discrete dispersion relation for that integrator. Thus, in the first part of this chapter, I will discuss the dispersion relation for a multi-Hamiltonian PDE, paying particular attention to the explicit ODEs obtained in Chapter 3 from a Lobatto IIIA–IIIB discretisation in space of an appropriate PDE. In the second part of this chapter, I consider the order of these explicit ODEs and how this order may be increased by an appropriate choice of initial conditions.

5.1 Dispersion Relations

A dispersion relation describes how the frequencies ω in time of a plane wave solution to a linearised PDE are related to the wave numbers ξ in space of that solution. While the behaviour of a nonlinear PDE will only be determined by the dispersion relation in regions where the linearised PDE is a valid approximation to the nonlinear PDE (such as near fixed points), it is often these regions that are of interest.

For numerical integrators a discrete dispersion relation may also be calculated which is an approximation to the continuous dispersion relation over a finite set of wave numbers. Properties of the numerical solution of PDEs, such as the stability, conservation of the sign of the wave group velocity, and the existence of spurious waves, may be determined by comparing the dispersion relation of the continuous PDE with that of the numerical integrator. See, for example, [2, 3, 15, 27].

The dispersion relation for a multi-Hamiltonian PDE (1.6) is calculated by first linearising the PDE to get

$$\boldsymbol{K}\boldsymbol{z}_t + \boldsymbol{L}\boldsymbol{z}_x = \boldsymbol{S}\boldsymbol{z}, \tag{5.1}$$

where $\boldsymbol{S}$ is a symmetric matrix such that $\boldsymbol{S}\boldsymbol{z}$ is the linear component of $\nabla_{\boldsymbol{z}} S(\boldsymbol{z})$.

Next, one assumes a periodic, plane wave solution to the PDE that is a pure exponential in t and x, i.e.,

$$\boldsymbol{z} = \mathrm{e}^{\mathrm{i}(\omega t + \xi x)} \tilde{\boldsymbol{z}}, \tag{5.2}$$

where $\tilde{\boldsymbol{z}}$ is a constant vector. Substituting Eq. (5.2) into Eq. (5.1) gives

$$(\mathrm{i}\omega \boldsymbol{K} + \mathrm{i}\xi \boldsymbol{L} - \boldsymbol{S})\, \tilde{\boldsymbol{z}} = \boldsymbol{0}, \tag{5.3}$$

which is satisfied when ω and κ satisfy the dispersion relation [8]

$$\det\left(\mathrm{i}\omega \boldsymbol{K} + \mathrm{i}\xi \boldsymbol{L} - \boldsymbol{S}\right) = 0. \tag{5.4}$$

In general, this dispersion relation is a polynomial of degree n (since $\boldsymbol{z} \in \mathbb{R}^n$) in both ω and κ, which can have different numbers of solutions, $\omega(\xi)$, depending on the value of ξ.

To obtain the discrete dispersion relation for a numerical method, the travelling wave solution is substituted into the discretised form of Eq. (5.1). This typically results in more complicated functions of ω and κ in Eq. (5.4), which may have a different number of solutions, $\omega(\xi)$, than the continuous dispersion relation for a given value of ξ.

As an example, consider the linearised wave equation, $u_{tt} = u_{xx} - \rho u$, which takes the form

$$\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_t \\ v_t \\ w_t \end{bmatrix} + \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_x \\ v_x \\ w_x \end{bmatrix} = \begin{bmatrix} \rho & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} \tag{5.5}$$

and has a continuous dispersion relation given by

$$\omega^2 = \xi^2 + \rho. \tag{5.6}$$

Discretising Eq. (5.5) with the Preissman box scheme, which is obtained by applying 1-stage Gaussian RK in space and time (and hence is multisymplectic by Theorem 3.2.1) and assuming a solution of the form

$$\boldsymbol{z} = \mathrm{e}^{\mathrm{i}(n\omega\Delta t + i\xi\Delta x)} \tilde{\boldsymbol{z}}, \tag{5.7}$$

gives the dispersion relation

$$\det\left(\mathrm{i}\frac{2}{\Delta t}\tan(\frac{1}{2}\omega\Delta t)\boldsymbol{K} + \mathrm{i}\frac{2}{\Delta x}\tan(\frac{1}{2}\xi\Delta x)\boldsymbol{L} - \boldsymbol{S}\right) = 0, \tag{5.8}$$

which can be written as

$$\left(\frac{2}{\Delta t}\tan(\frac{1}{2}\omega\Delta t)\right)^2 = \left(\frac{2}{\Delta x}\tan(\frac{1}{2}\kappa\Delta x)\right)^2 + \rho. \tag{5.9}$$

The Preissman box scheme has received a lot of attention in the past due to its ability to avoid developing instabilities in the numerical solution, a feature indicative of parasitic waves. This property has been attributed to the qualitative preservation of the dispersion relation, i.e., the continuous set of frequencies on the positive real line are mapped diffeomorphically onto the domain $(-\pi, \pi)$ in both space and time.

On the other hand, discretising Eq. (5.5) in space with 3-stage Lobatto IIIA–IIIB and assuming a solution of the form

$$\boldsymbol{Z}_{i,j} = \mathrm{e}^{\mathrm{i}(\omega t + (i+c_j)\xi\Delta x)} \widetilde{\boldsymbol{Z}}_{0,j} \tag{5.10}$$

for each of the stage variables $\boldsymbol{Z}_{i,j}$ gives the following dispersion relation

$$\det\left(\begin{bmatrix} (\Delta x)^2(\omega^2 - \rho) - 14 - 2\cos(\xi\Delta x) & 16\cos(\frac{1}{2}\xi\Delta x) \\ 8\cos(\frac{1}{2}\xi\Delta x) & (\Delta x)^2(\omega^2-\rho) - 8 \end{bmatrix}\right) = 0 \tag{5.11}$$

after eliminating the $V_{i,j}$ and $W_{i,j}$ variables. Solving this dispersion relation for $\omega\Delta x$ gives

$$\omega\Delta x = \pm\left(\rho(\Delta x)^2 + 11 + \cos(\xi\Delta x) \pm \left(73 + 70\cos(\xi\Delta x) + \cos^2(\xi\Delta x)\right)^{\frac{1}{2}}\right)^{\frac{1}{2}}. \tag{5.12}$$

For each value of ξ, Eq. (5.12) admits four values of ω, two of which closely approximate the value of ω in the continuous dispersion relation for small values of ξ. A comparison of the discrete and continuous dispersion relations from Eqs. (5.12) and (5.6) respectively is shown in Figure 5.1.

Generally, such extra solutions to the discrete dispersion relation indicates that the numerical integrator supports spurious modes, which appear as unwanted parasitic waves in the numerical solution to the PDE. However, as will be shown below, for RK and PRK methods applied to multi-Hamiltonian PDEs that satisfy the conditions of Theorem 3.4.1, these extra solutions to the discrete dispersion relation have relevance and do not correspond to parasitic waves.

5.1.1 Stability

It was noted in Section 3.4.2 that for the class of multi-Hamiltonian PDEs satisfying the conditions of Theorem 3.4.1, the variables $\boldsymbol{p}$ can be eliminated and the multi-Hamiltonian PDE can be written in the form

$$\boldsymbol{q}_{xx} = f(\boldsymbol{q}, \boldsymbol{q}_t), \tag{5.13}$$

where $\boldsymbol{q} \in \mathbb{R}^{\mathrm{rank}(\boldsymbol{K})}$. Now, if a periodic plane wave solution of the form of Eq. (5.2) is assumed, then $\boldsymbol{q}_t = \mathrm{i}\omega\boldsymbol{q}$ and discretising Eq. (5.13) in space is equivalent to finding periodic solutions of the ODEs

$$\boldsymbol{q}_{xx} = f(\boldsymbol{q}). \tag{5.14}$$

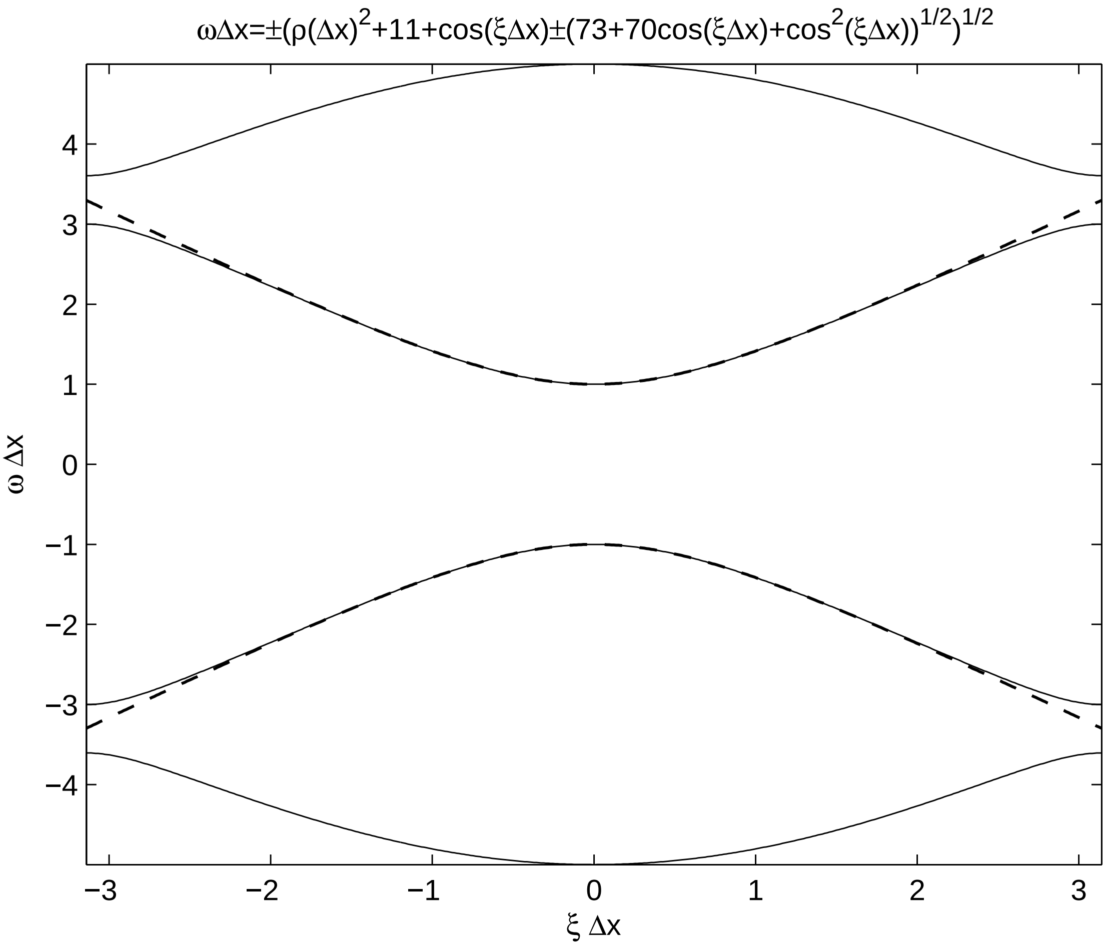


Figure 5.1: A comparison between the discrete dispersion relation (solid line) for $\rho(\Delta x)^2 = 1$ and the continuous dispersion relation (dashed line) for $\rho = 1$.

Furthermore, if $f(\boldsymbol{q})$ is linear then Eq. (5.14) can be decoupled into a set of harmonic oscillators with frequencies ω^γ for $\gamma = 1, \ldots, \mathrm{rank}(\boldsymbol{K})$. This gives a dispersion relation,

$$-\xi^2 = \omega^\gamma, \tag{5.15}$$

for each harmonic oscillator subsystem labelled by γ, while the discrete dispersion relation for each harmonic oscillator subsystem is that of the PRK discretisation applied to a harmonic oscillator of frequency ω^γ.

Suppose that such a PRK discretisation applied to a harmonic oscillator of frequency ω^γ gives a linear map

$$q^\gamma \mapsto R(\Delta x \omega^\gamma) q^\gamma. \tag{5.16}$$

If the coefficients of the RK or PRK method satisfies the symplecticity condition (Eq. (3.6)), then R has determinant 1. Furthermore, if the trace of R is at most 2 in absolute value, then the map (5.16) is conjugate to a rotation by angle ξ where $\operatorname{tr}(R) = 2\cos(\xi)$. ($R$ is known as the stability function and, for an RK method, takes the form [19]

$$R(z) = 1 + z\boldsymbol{b}^T(\boldsymbol{I} - z\boldsymbol{A})^{-1}\mathbf{1}. \tag{5.17}$$

For Gaussian RK methods, it is an approximation to e^z with an error of $\mathcal{O}((\Delta x)^{2r+1})$.)

The stability of such a RK or PRK discretisation as a time integrator can be determined as the largest value of the time step Δt such that $|\operatorname{tr}(R)| \leq 2$, i.e., the largest value of Δt such that the map (5.16) remains conjugate to a rotation. Similarly, the stability of such a RK or PRK discretisation in space can also be determined from $\operatorname{tr}(R)$. For a given wave number, ξ, in space, the values of ω^γ such that $\operatorname{tr}(R(\Delta x\omega^\gamma)) = 2\cos(\xi)$ determine the dispersion relation. If the RK or PRK method has r stages at r distinct quadrature points and the modes of each harmonic oscillator subsystem are to be periodic, then there must be r distinct values of ω. If there are fewer than r distinct values of ω then the method will, in general, be unstable.

For an r-stage Gaussian RK discretisation in space there are precisely r modes for each value of ξ, thus, if the method is well defined, it will be stable. Furthermore, for these methods the stability function is invertible, in general, and $|\operatorname{tr}(R)| \leq 2$ for all values of ω^γ. Therefore, for this class of multi-Hamiltonian PDEs, the discrete dispersion relation is conjugate to the continuous dispersion relation. Moreover, it is monotonic and continuous.

For an r-stage Lobatto IIIA–IIIB PRK discretisation in space there are precisely $r-1$ modes for each value of ξ (as is hinted at in Figure 5.1, the dispersion relation for the 3-stage method has twice as many modes as the continuous dispersion relation), however the last mode is not unstable as the last quadrature point coincides with the first quadrature point of the next cell and there is effectively only $r-1$ active variables per cell. Thus the explicit ODEs formed by applying a Lobatto IIIA–IIIB discretisation in space to a multi-Hamiltonian PDE satisfying the conditions of Theorem 3.4.1 is a stable discretisation.

Furthermore, the regions where $|\operatorname{tr}(R)| > 2$ do not correspond to instability (as they would do in a time discretisation), but rather to jumps in the value of ω in the dispersion relation. Therefore, for this class of multi-Hamiltonian PDEs, the discrete dispersion relation is conjugate to a portion of the continuous dispersion relation. This conjugacy is monotonic but no longer necessarily continuous, as it was for Gaussian RK discretisations.

A numerical demonstration of the stability of the Lobatto IIIA–IIIB discretisation in space is given in Figures 5.2, 5.3 and 5.4. To obtain these figures, the NLS equation is discretised in space with 3-stage Lobatto IIIA–IIIB and in time with 2-stage Lobatto IIIA–IIIB. The spatial domain is periodic on $[0, L]$, where $L = 4\sqrt{(2)}\pi$, and divided into 256 equally spaced nodes. The initial conditions are given by $P_{i,1} = f(i\Delta x)$, $P_{i,2} = f((i+\frac{1}{2})\Delta x)$ and $Q_{i,1} = Q_{i,2} = 0$, where $f(x) = 1.5(1 - 0.1\cos(2\pi x/L))$. This integrator is then stepped forwards in time for 10^7 steps with a step size of 10^{-4}. (This step size

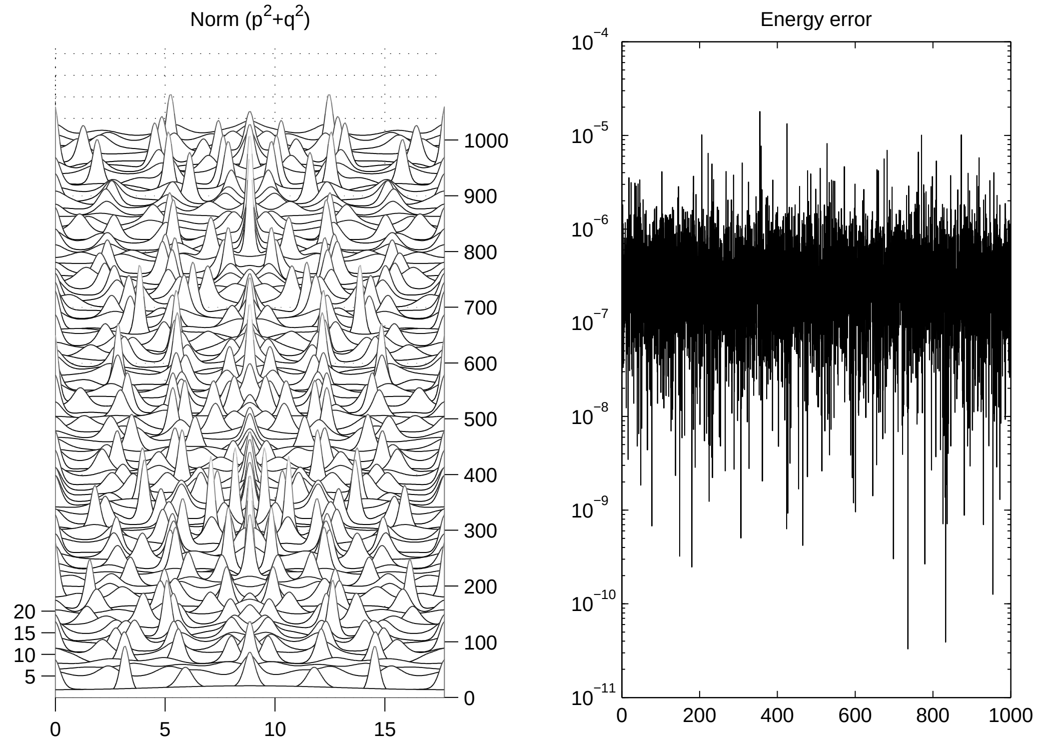


Figure 5.2: A waterfall plot of the norm $(p^2 + q^2)$ and the energy error for the NLS equation.

is chosen such that the quadratic equations in Eq. (4.12) can be solved. For many discretisations in time, the step size will also be restricted by $\Delta t < C(\Delta x)^2$ in order for the integrator to remain stable.)

In Figure 5.2, a waterfall plot of the norm $(p^2 + q^2)$ is shown on the left, while a plot of the error in the total energy is shown on the right. The energy behaviour is what one would expect from a symplectic integrator, that is, the total energy is approximately conserved by the integrator. Remarkably, for symmetric initial conditions this integrator appears to preserve the symmetry of the solution exactly. In figure 5.3, the solution (p and q) at $t = 1000$ (i.e., after 10^7 steps) is given. This solution appears to be smooth, i.e., there appears to be no high frequency oscillations in space, which is an indication of the stability of the integrator. The Fourier transform of p at $t = 1000$ in Figure 5.4 (the Fourier transform of q is almost identical) confirms the lack of any high frequency oscillations in space.

5.2 Order

In [33], it is shown that a RK or PRK method is equivalent to a collocation or discontinuous collocation method where the node and stage values of the RK or PRK method match

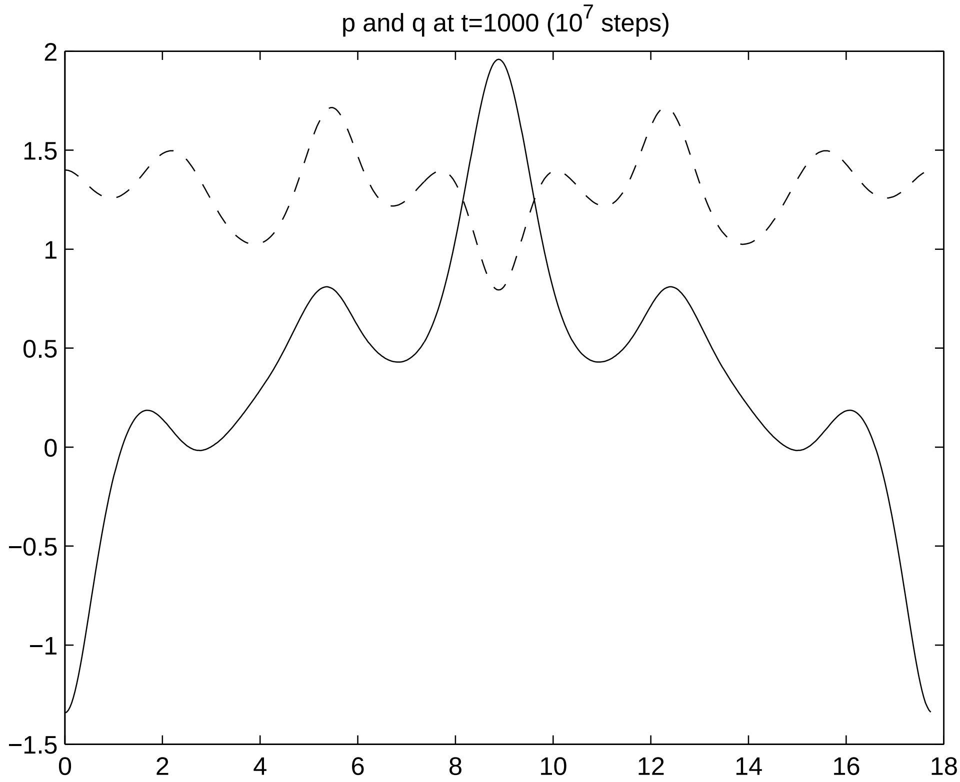


Figure 5.3: The values of p (solid line) and q (dashed line) after 10^7 steps of size 10^{-4} showing a lack of any high frequency wiggles in the solution.

the collocation polynomial at the quadrature points. Moreover, by a series of theorems and lemmas, the order of these methods and their collocation polynomials are determined. Thus, the order of the node and stage values of the RK or PRK method are determined by the order of the collocation polynomial at the quadrature points. For the convenience of the reader, I have given a direct quote of these theorems and lemmas (with only adjustments to the equation numbers for readability) in Appendix B.

These theorems and lemmas show that the order of a collocation or discontinuous collocation method is given by the largest value of ξ such that $B(\xi)$ holds, and that if $B(\xi)$ holds for some $\xi > r$, then the method has the same order as the underlying quadrature formula. They also show that the order of the collocation polynomial is given by the largest value of ξ such that $C(\xi)$ holds, which is $\xi = r$ for an r-stage collocation method and $\xi = r - 2$ for a discontinuous collocation method.

These conditions ($B(\xi)$ and $C(\xi)$) for determining the order of a RK or PRK method can also be seen by considering Taylor expansions of the collocation polynomial around the first node of the RK or PRK method. Consider an ODE, $\partial_x y(x) = f(x, y)$, discretised by an r-stage RK or PRK method. Let $u(x)$ be the collocation polynomial of the equivalent

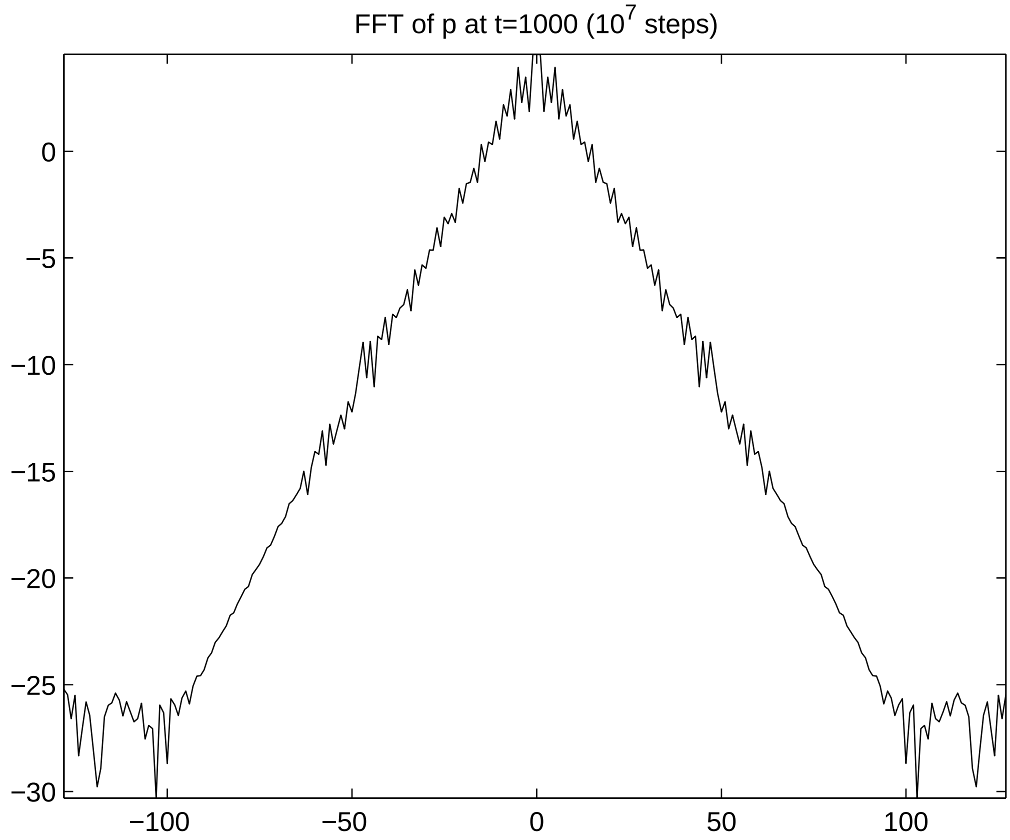


Figure 5.4: The log of the fast Fourier transform of p after 10^7 steps of size 10^{-4} showing that the high frequency components are exponentially small. The log of the fast Fourier transform of q after 10^7 steps is almost identical.

collocation method where

$$\begin{aligned} u(0) &= y(0) = y_0, \\ \partial_x u(c_i \Delta x) &= f(c_i \Delta x, u(c_i \Delta x)). \end{aligned} \tag{5.18}$$

Now, a Taylor expansion of the actual solution at $x = c_i \Delta x$ is given by

$$y(c_i \Delta x) = \sum_{k=0}^{\infty} \frac{(\Delta x)^k}{k!} y^{(k)}(0), \tag{5.19}$$

where $y^{(k)}(x)$ is the kth derivative of $y(x)$. Similarly, a Taylor expansion of the function $f(c_i \Delta x, u(c_i \Delta x))$ about $x = 0$ is given by

$$\begin{aligned} f(c_i \Delta x, u(c_i \Delta x)) &= \sum_{k=0}^{\infty} \frac{(\Delta x)^k}{k!} f^{(k)}(0, u(0)) \\ &= \sum_{k=1}^{\infty} \frac{(\Delta x)^{k-1}}{(k-1)!} y^{(k)}(0). \end{aligned} \tag{5.20}$$

Therefore, a Taylor expansion of the collocation polynomial for the RK or PRK method at $x = \Delta x$ (i.e., the node y_1 of the RK or PRK method) is given by

$$\begin{aligned} y_1 = u(\Delta x) &= u(0) + \Delta x \sum_{i=1}^{r} b_i f(c_i \Delta x, u(c_i \Delta x)) \\ &= y(0) + \Delta x \sum_{i=1}^{r} b_i \sum_{k=1}^{\infty} \frac{(\Delta x)^{k-1}}{(k-1)!} y^{(k)}(0) \\ &= y(0) + \sum_{k=1}^{\infty} k \sum_{i=1}^{r} b_i c_i^{k-1} \frac{(\Delta x)^k}{k!} y^{(k)}(0) \\ &= y(\Delta x) + (O)((\Delta x)^{\xi+1}), \end{aligned} \tag{5.21}$$

where ξ is the largest integer such that $B(\xi)$ is satisfied.

Similarly, a Taylor expansion of the collocation polynomial for the RK or PRK method at the quadrature point $c_i \Delta x$ (i.e., the stage $Y_{0,i}$ of the RK or PRK method) is given by

$$\begin{aligned} Y_{0,i} = u(c_i \Delta x) &= u(0) + \Delta x \sum_{j=1}^{r} a_{ij} f(c_j \Delta x, u(c_j \Delta x)) \\ &= y(0) + \Delta x \sum_{j=1}^{r} a_{ij} \sum_{k=1}^{\infty} \frac{(c_j \Delta x)^{k-1}}{(k-1)!} y^{(k)}(0) \\ &= y(0) + \sum_{k=1}^{\infty} k \sum_{j=1}^{r} a_{ij} c_j^{k-1} \frac{(\Delta x)^k}{k!} y^{(k)}(0) \\ &= y(c_i \Delta x) + (O)((\Delta x)^{\xi+1}), \end{aligned} \tag{5.22}$$

where ξ is the largest integer such that $C(\xi)$ is satisfied.

If an ODE is integrated over the domain $[0, L]$ then the total error incurred by the RK or PRK method is of the order of the error per step, $\mathcal{O}((\Delta x)^{\xi+1})$, multiplied by the number of steps, $L/\Delta x$, which gives the order of the RK or PRK method as $\mathcal{O}((\Delta x)^{\xi})$.

As has been mentioned earlier, for multi-Hamiltonian PDEs discretised by RK or PRK methods in space, it is not the nodes but rather the stage variables that are the active variables. Thus, for a RK or PRK discretisation in space of a multi-Hamiltonian PDE it is the order of the stage variables that is important, not the order of the nodes. Furthermore, the initial conditions for such a discretisation are typically chosen to exactly match the initial conditions of the PDE at the nodes, so the above accumulation of the errors does not occur. Thus the order of the stage variables in the spatial discretisation is $\xi+1$ instead of ξ, where ξ is the largest integer such that $C(\xi)$ is satisfied.

For r-stage RK methods based on collocation, $B(\xi)$ and $C(\xi)$ are satisfied for $\xi \leq r$ (since that is how they are constructed) but not $\xi = r+1$, in general, therefore the stage variables in these methods are all at least of order $r+1$. Whereas r-stage RK methods based on discontinuous collocation (which are constructed from $B(\xi)$ and $D(\xi)$ for $\xi = r$)

are equivalent to $r-2$ stage RK methods based on collocation and thus the stage variables are at least of order $r-1$.

An important consequence of this observation is that the multisymplectic integrators considered in [27, 38, 63], which are formed by discretising a multi-Hamiltonian PDE in time and space with RK or PRK methods, are (when they are well defined) not of the order of the RK or PRK method (e.g., $2r$ for GRK methods), but of the order of the stage variables instead (e.g., $r+1$ for GRK methods).

The Lobatto IIIA–IIIB class of PRK methods, which are of particular interest in this thesis, can be considered to be a discontinuous collocation method (as given in the proof of Theorem 2.2 of [33], which is quoted in Appendix B). The stage variables of an r-stage Lobatto IIIA–IIIB discretisation in space of a PDE are at least of order r for the variables that Lobatto IIIA is applied to and at least of order $r-1$ for the variables that Lobatto IIIB is applied to, which can be seen from the order of the collocation polynomial in Eq. (B.28).

An important point to consider when determining the order of a discretisation of a PDE in space by a RK or PRK method is that the ODEs one obtains (whether explicit or implicit) contain an approximation of the spatial derivatives in the PDE. The order of this approximation of the spatial derivatives may be less than the order of the stage variables, in which case, the order of the discretisation in space of the PDE is the minimum of the order of each of the spatial derivatives.

For example, for a multi-Hamiltonian PDE satisfying the conditions of Theorem 3.4.1, the ODEs that one obtains from discretising in space with an r-stage Lobatto IIIA–IIIB method contain an approximation to the second derivative in space of the $\boldsymbol{q}$ variables. Direct calculation of the Taylor expansion of this approximation, for $r \leq 10$, shows that this approximation to the second derivative is of order $r-1$ at each of the quadrature points (an example for $r=3$ is given in the next section) and of order r at the first quadrature point when r is even (due to the symmetry of the approximation). Thus we have the following conjecture:

Conjecture 5.2.1. *The explicit ODEs obtained by applying an r-stage Lobatto IIIA–IIIB discretisation in space to a multi-Hamiltonian PDE satisfying the conditions of Theorem 3.4.1 are an approximation to the PDE of order $r-1$. For $r=2$, the ODEs are an approximation to the PDE of order* 2.

In the following section, I show how the order of these ODEs (for $r>2$) may be increased by an appropriate choice of the initial conditions.

5.2.1 Initial Conditions

Generally, when one specifies a set of initial conditions for a PDE that has been discretised on a grid, one specifies those initial conditions on the variables located at the nodes of the grid. However, as mentioned in the previous section, it is the stage variables that are the active variables in the spatial discretisation (except for when $r = 2$ where the node variables and the stage variables coincide). Therefore, it is necessary to specify the initial conditions on the stage variables, not the node variables. But this assignment of the initial conditions such that it matches the initial conditions of the continuous PDE at the nodes of the grid is not unique for $r > 2$ and furthermore, may affect the order of the numerical integrator.

For example, given the initial condition $z^\gamma(x) = f(x)$ for a continuous PDE, a naïve way of assigning this initial condition to the stage variables of the discretised system is to simply let $Z^\gamma_{i,j} = f((i + c_j)\Delta x)$ for all stages j. The accuracy of the approximation to the second derivative, z^γ_{xx}, in the discrete ODEs is then found by considering Taylor expansions of the initial conditions about each stage variable.

For a discretisation with $r = 3$, a Taylor expansion of the initial conditions about the first stage variable gives

$$\begin{aligned}
\partial_x^2 Z^\gamma_{i,1} &= -\frac{1}{(\Delta x)^2}\Big(f - \Delta x f^{(1)} + \frac{(\Delta x)^2}{2} f^{(2)} - \frac{(\Delta x)^3}{6} f^{(3)} + \frac{(\Delta x)^4}{24} f^{(4)} - \cdots\Big) \\
&\quad + \frac{8}{(\Delta x)^2}\Big(f - \frac{\Delta x}{2} f^{(1)} + \frac{(\Delta x)^2}{8} f^{(2)} - \frac{(\Delta x)^3}{48} f^{(3)} + \frac{(\Delta x)^4}{384} f^{(4)} - \cdots\Big) \\
&\quad - \frac{14}{(\Delta x)^2} f \\
&\quad + \frac{8}{(\Delta x)^2}\Big(f + \frac{\Delta x}{2} f^{(1)} + \frac{(\Delta x)^2}{8} f^{(2)} + \frac{(\Delta x)^3}{48} f^{(3)} + \frac{(\Delta x)^4}{384} f^{(4)} + \cdots\Big) \\
&\quad - \frac{1}{(\Delta x)^2}\Big(f + \Delta x f^{(1)} + \frac{(\Delta x)^2}{2} f^{(2)} + \frac{(\Delta x)^3}{6} f^{(3)} + \frac{(\Delta x)^4}{24} f^{(4)} + \cdots\Big) \\
&= f^{(2)} + \mathcal{O}((\Delta x)^2),
\end{aligned} \tag{5.23}$$

where $f^{(n)}$ is the nth derivative of $f(x)$ evaluated at $x = i\Delta x$.

Similarly, a Taylor expansion of the initial conditions about the second stage variable gives

$$\begin{aligned}
\partial_x^2 Z^\gamma_{i,2} &= \frac{4}{(\Delta x)^2}\Big(f - \frac{\Delta x}{2} f^{(1)} + \frac{(\Delta x)^2}{8} f^{(2)} - \frac{(\Delta x)^3}{48} f^{(3)} + \frac{(\Delta x)^4}{384} f^{(4)} - \cdots\Big) \\
&\quad - \frac{8}{(\Delta x)^2} f \\
&\quad + \frac{4}{(\Delta x)^2}\Big(f + \frac{\Delta x}{2} f^{(1)} + \frac{(\Delta x)^2}{8} f^{(2)} + \frac{(\Delta x)^3}{48} f^{(3)} + \frac{(\Delta x)^4}{384} f^{(4)} + \cdots\Big) \\
&= f^{(2)} + \mathcal{O}((\Delta x)^2),
\end{aligned} \tag{5.24}$$

where $f^{(n)}$ is now evaluated at $(i + \frac{1}{2})\Delta x$. Thus, with this naïve assignment of the initial conditions, the approximation to z^γ_{xx} for $r = 3$ is only second order, in accordance with Conjecture 5.2.1.

This order may be sufficient for some applications, but a higher order discretisation in space is desirable in general. Since the stage variables at the ends of a cell are equal to the node variables, it is desirable that the first and last stage variables be specified as if they were node variables, i.e., unmodified. However, it may be possible to specify the initial conditions on the other stage variables as a modified function of the initial conditions for the continuous PDE such that the order of the approximation of the second derivatives in space is increased.

For example, for $r = 3$, the following modification to the initial conditions on the second stage variable increases the order of the spatial discretisation from second order to fourth order. Let the initial conditions be specified as

$$\begin{aligned} Z^{\gamma}_{i,1} &= f(i\Delta x), \\ Z^{\gamma}_{i,2} &= \tilde{f}((i+\frac{1}{2})\Delta x) = f((i+\frac{1}{2})\Delta x) + \frac{(\Delta x)^4}{384} f^{(4)}((i+\frac{1}{2})\Delta x), \\ Z^{\gamma}_{i,3} &= Z^{\gamma}_{i+1,1} = f((i+1)\Delta x). \end{aligned} \tag{5.25}$$

Then Eq. (5.23) becomes

$$\begin{aligned} \partial_x^2 Z^{\gamma}_{i,1} = &-\frac{1}{(\Delta x)^2}\Big(f - \Delta x f^{(1)} + \frac{(\Delta x)^2}{2} f^{(2)} - \frac{(\Delta x)^3}{6} f^{(3)} + \frac{(\Delta x)^4}{24} f^{(4)} - \cdots\Big) \\ &+ \frac{8}{(\Delta x)^2}\Big(\tilde{f} - \frac{\Delta x}{2}\tilde{f}^{(1)} + \frac{(\Delta x)^2}{8}\tilde{f}^{(2)} - \frac{(\Delta x)^3}{48}\tilde{f}^{(3)} + \frac{(\Delta x)^4}{384}\tilde{f}^{(4)} - \cdots\Big) \\ &- \frac{14}{(\Delta x)^2} f \\ &+ \frac{8}{(\Delta x)^2}\Big(\tilde{f} + \frac{\Delta x}{2}\tilde{f}^{(1)} + \frac{(\Delta x)^2}{8}\tilde{f}^{(2)} + \frac{(\Delta x)^3}{48}\tilde{f}^{(3)} + \frac{(\Delta x)^4}{384}\tilde{f}^{(4)} + \cdots\Big) \\ &- \frac{1}{(\Delta x)^2}\Big(f + \Delta x f^{(1)} + \frac{(\Delta x)^2}{2} f^{(2)} + \frac{(\Delta x)^3}{6} f^{(3)} + \frac{(\Delta x)^4}{24} f^{(4)} + \cdots\Big) \\ = &\, f^{(2)} + \mathcal{O}((\Delta x)^4) \end{aligned} \tag{5.26}$$

where $f^{(n)}$ is evaluated at $i\Delta x$.

Similarly, Eq. (5.24) becomes

$$\begin{aligned} \partial_x^2 Z^{\gamma}_{i,2} = &\frac{4}{(\Delta x)^2}\Big(f - \frac{\Delta x}{2} f^{(1)} + \frac{(\Delta x)^2}{8} f^{(2)} - \frac{(\Delta x)^3}{48} f^{(3)} + \frac{(\Delta x)^4}{384} f^{(4)} - \cdots\Big) \\ &- \frac{8}{(\Delta x)^2}\tilde{f} \\ &+ \frac{4}{(\Delta x)^2}\Big(f + \frac{\Delta x}{2} f^{(1)} + \frac{(\Delta x)^2}{8} f^{(2)} + \frac{(\Delta x)^3}{48} f^{(3)} + \frac{(\Delta x)^4}{384} f^{(4)} + \cdots\Big) \\ = &\, f^{(2)} + \mathcal{O}((\Delta x)^4) \\ = &\, \tilde{f}^{(2)} + \mathcal{O}((\Delta x)^4) \end{aligned} \tag{5.27}$$

where $f^{(n)}$ and $\tilde{f}^{(n)}$ are evaluated at $(i + \frac{1}{2})\Delta x$.

To numerically demonstrate this increase in order, consider the sine-Gordon wave equation

$$u_{tt} = u_{xx} - \sin(u), \tag{5.28}$$

which is the nonlinear wave equation with the potential $V(u) = -\cos(u)$. This is discretised in space with 3-stage Lobatto IIIA–IIIB to give the explicit ODEs in Eq. (3.56), which are then discretised in time by 2-stage Lobatto IIIA–IIIB to create an explicit multisymplectic integrator. The initial conditions are given by

$$\begin{aligned} U_{i,1}(0) &= f(i\Delta x), \\ U_{i,2}(0) &= f((i+\frac{1}{2})\Delta x), \\ \partial_t U_{i,1}(0) &= \partial_t U_{i,2}(0) = 0, \end{aligned} \tag{5.29}$$

where $f(x) = \exp(\cos(2\pi x/L))$ and $L = 4\sqrt{2}\pi$, and one step forwards in time, of size $\Delta t = 10^{-2}$, is taken for the three cases, $\Delta x = L/64$, $\Delta x = L/128$ and $\Delta x = L/256$.

The order of the method can then be found by calculating

$$\log_2\left(\frac{|U_{64}(\Delta t) - U_{128}(\Delta t)|}{|U_{128}(\Delta t) - U_{256}(\Delta t)|}\right), \tag{5.30}$$

where U_{64} is the solution for $\Delta x = L/64$, etc. For these unmodified initial conditions, the order is shown in Figure 5.5, while for the modified initial conditions ($U_{i,2}(0) = f((i+\frac{1}{2})\Delta x) + \frac{(\Delta x)^4}{384} f^{(4)}((i+\frac{1}{2})\Delta x)$) the order is shown in Figure 5.6.

For higher values of r, the appropriate modification to the initial conditions for the inner stage variables is not yet known. However, it is expected that an expression for the appropriate modification to the initial conditions does exist and that this will, in general, increase the order in space of an integrator formed from such a system of ODEs by 1, i.e., to the order of the stage variables in the spatial discretisation.

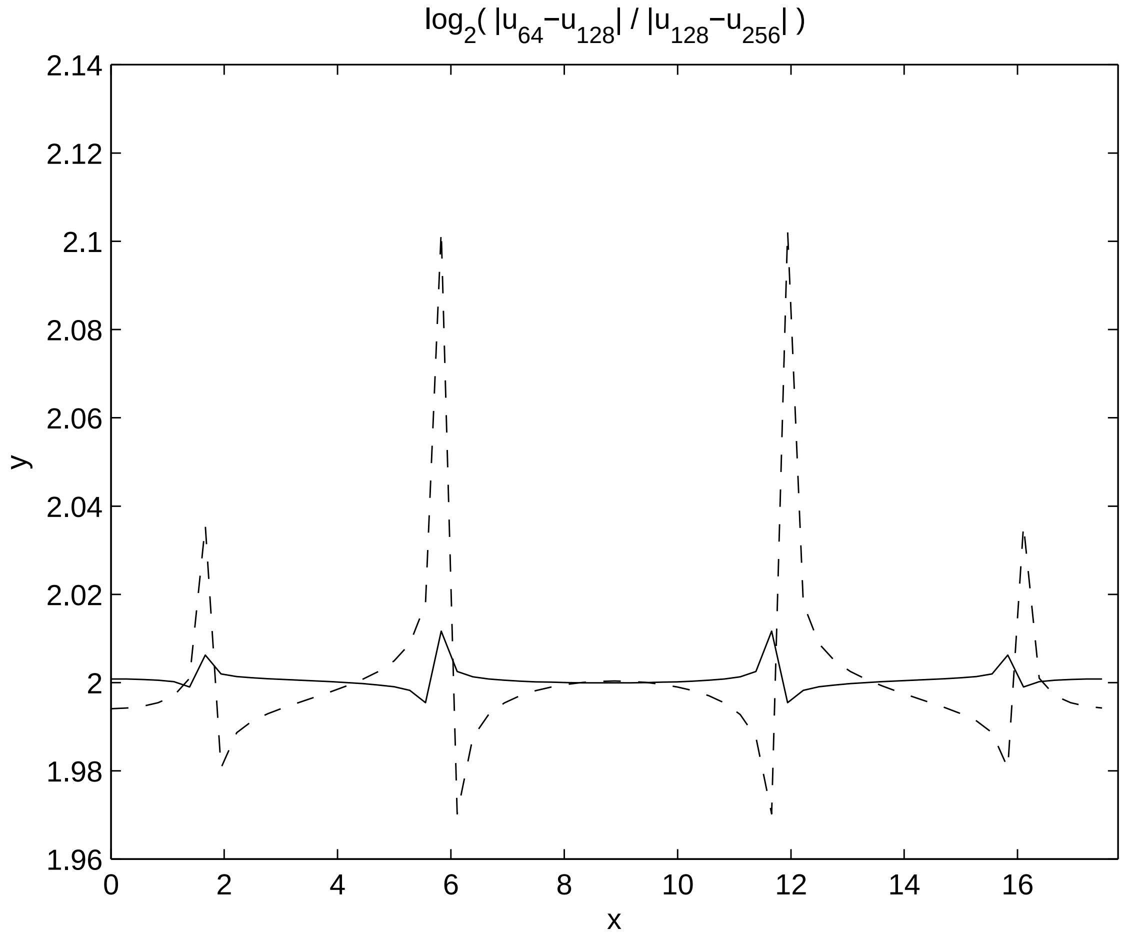


Figure 5.5: The integrator for the sine–Gordon equation given by 3-stage Lobatto IIIA–IIIB discretisation in space and 2-stage Lobatto IIIA–IIIB discretisation in time, with unmodified initial conditions, has order 2. The dashed line gives the order at $i\Delta x$, the solid line gives the order at $(i + \frac{1}{2})\Delta x$.

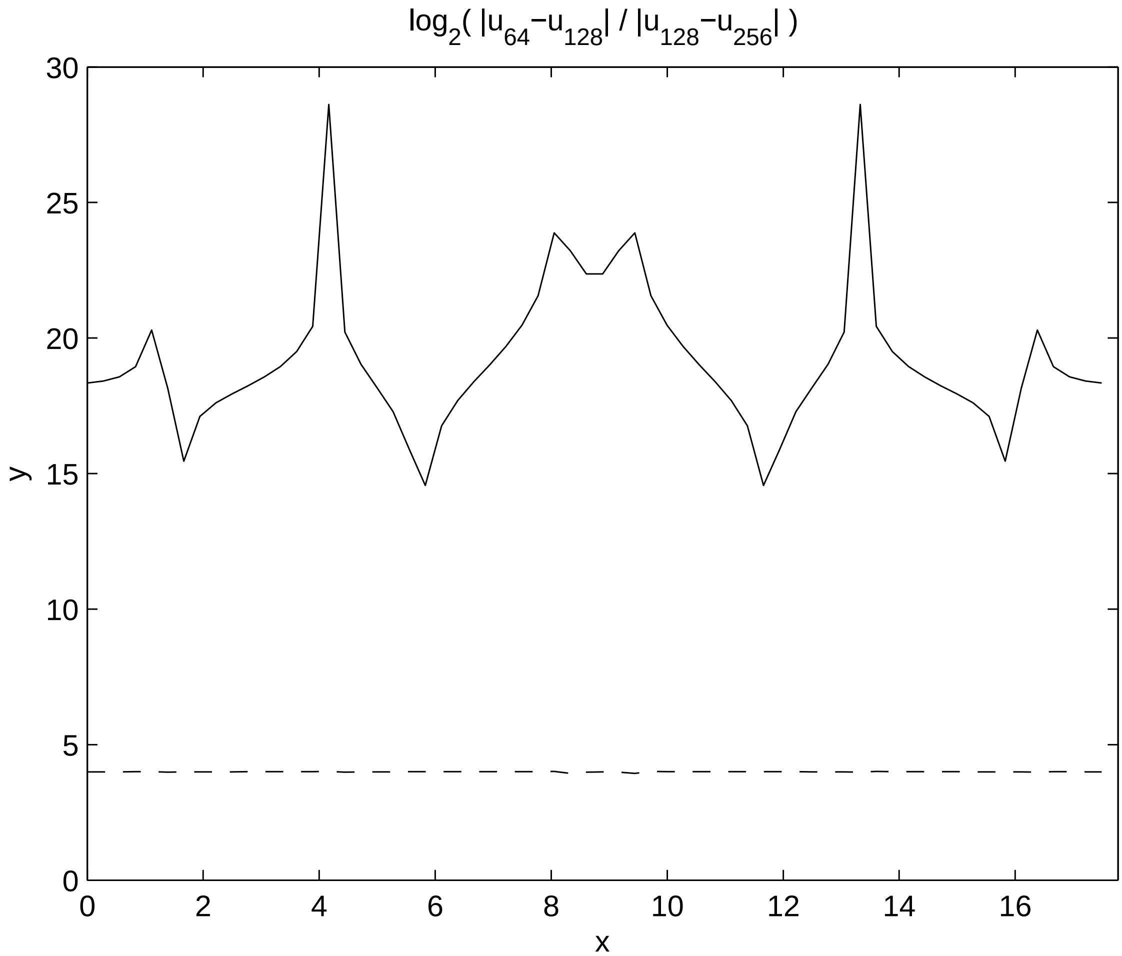


Figure 5.6: The integrator for the sine–Gordon equation given by 3-stage Lobatto IIIA–IIIB discretisation in space and 2-stage Lobatto IIIA–IIIB discretisation in time, with modified initial conditions, has order 4. The dashed line gives the order at $i\Delta x$, the solid line gives the order at $(i + \frac{1}{2})\Delta x$.
Due to the small value of Δt, the order of the stage 2 variables appear unusually high since $|U_{128} - U_{256}|$ is near round-off error. As Δt is increased, the order of the stage 2 variables reduces to 4, but the integrator becomes unstable due to $\Delta t >> \Delta x$.

Chapter 6

Conclusion

6.1 Summary and closing remarks

Throughout this thesis, I have considered the topic of multisymplectic integration by numerical integrators that preserve a discrete analogue of a multisymplectic conservation law, otherwise known as multisymplectic integrators.

Much of the focus of this thesis has been on multisymplectic integrators constructed from Runge–Kutta (RK) or partitioned Runge–Kutta (PRK) discretisations of a multi-Hamiltonian PDE. I have shown that, for a general PRK discretisation (which includes RK discretisations) of a general multi-Hamiltonian PDE, the resulting system of equations will satisfy a natural discrete analogue of the continuous multisymplectic conservation law associated with that multi-Hamiltonian PDE when the coefficients of the PRK method satisfy a simple set of conditions. However, simply satisfying a discrete multisymplectic conservation law is not sufficient for a system of equations to be labelled as a multisymplectic integrator; the system of equations must also form a well-defined numerical method and remain well defined when boundary conditions are applied. Thus, a question that has driven much of this research, and has been an underlying theme of Chapters 2, 3 and 4, is whether or not such discretisations give rise to well-defined multisymplectic integrators.

In Chapter 2, I have shown that discretisation by Gaussian RK methods gives rise, in general, to a system of equations that is not well defined as a numerical method. In particular, for periodic boundary conditions, Theorem 2.2.1 gives the requirements for a semi-discretisation of a multi-Hamiltonian PDE by a Gaussian RK method to be well defined as being that both the number of stages in the method and the number of cells in the grid must be odd. For other types of boundary conditions, I have argued, based on the collocation polynomials that interpolate each cell of the semi-discretisation, that the spatial discretisation is not well defined in general. For a full discretisation of a multi-Hamiltonian PDE by Gaussian RK methods, the question of whether the discretisation is well defined depends on how the boundary conditions are applied. For many choices of how to apply the boundary conditions, the numerical integrator does not give an accurate representation of the PDE, while for other choices of boundary conditions, the numerical integrator is highly implicit, ill-conditioned and of little use, except in the simplest of applications. One exception to this is the Preissman box scheme, which is the lowest order multisymplectic integrator that can be constructed from Gaussian RK discretisation in time and space. The simplicity of the Preissman box scheme allows many of the difficulties associated with higher order Gaussian RK discretisations to be avoided.

In contrast to the ill-defined methods formed in Chapter 2 by RK discretisation, in Chapter 3, I have defined a class of multi-Hamiltonian PDEs that, when discretised in space by a method from the Lobatto IIIA–IIIB class of PRK methods, give rise to a system of explicit (and hence, well-defined) ODEs in time that are local, of high order in space and satisfy a local semi-discrete multisymplectic conservation law. Moreover, these explicit ODEs handle various types of boundary conditions in a simple (in some cases, trivial) and local manner, without the need for any further restrictions, as is often the case for other discretisations (e.g., the implicit midpoint method with periodic boundary conditions requires an odd number of grid points). This is a great boon for the implementation of these integrators, as boundary conditions can be applied at a boundary without the need to consider how they might affect other boundaries. This class of multi-Hamiltonian PDEs includes such famous equations as the nonlinear Schrödinger (NLS) equation, the nonlinear wave equation (including the Klein–Gordon and sine–Gordon equations) and the Boussinesq equation. Furthermore, I have given a construction algorithm that allows one to construct these explicit ODEs from the multi-Hamiltonian form of a PDE in this class.

Explicit (and hence, well-defined) multisymplectic integrators that are of high order in space may be formed from these explicit ODEs by applying an appropriate explicit discretisation in time. In Chapter 4, I have shown that for some PDEs, such as the nonlinear wave equation and the Boussinesq equation, applying a 2-stage Lobatto IIIA–IIIB discretisation in time to the explicit ODEs gives an explicit multisymplectic integrator that is local and of high order in space. Furthermore, the order in time of such integrators may be increased by composition. For other PDEs, such a discretisation in time is not explicit in general due to non-separable terms in the Hamiltonian. Nevertheless, if the nonlinear terms in the explicit ODEs resulting from these non-separable terms in the Hamiltonian appear quadratically, as is the case for the NLS equation, then such a discretisation may still be explicit. However, it will require solving scalar quadratic equations, which may impose further restrictions on the step size used in the time direction. I have used the NLS equation discretised thus as the first of three examples demonstrating how an explicit multisymplectic integrator may be formed from the explicit ODEs obtained in Chapter 3. The second and third examples of time discretisations are splitting methods for the NLS equation, the former being the standard linear–nonlinear splitting and the latter being a real–imaginary–nonlinear splitting. The multisymplectic integrators in the first and third examples are shown to satisfy a local multisymplectic conservation law, while the second is shown to satisfy a non-local multisymplectic conservation law. Interestingly, only the second method preserves the norm conservation law that is possessed by the NLS equation.

In Chapter 5, I have considered the stability of RK and PRK discretisations in space by analysing the dispersion relation for such discretisations. Of importance is that, for Gaussian RK methods, the dispersion relation exhibits the same number of solutions for the frequency as a function of the wave number as the number of stages in the Gaussian RK method. Similarly, for Lobatto IIIA–IIIB methods, the dispersion relation exhibits one less solution for the frequency as a function of wave number than the number of stages in the Lobatto IIIA–IIIB method. Generally, this is an indication that the numerical integrator supports spurious modes, which tend to destabilise the solution. However, for Gaussian RK and Lobatto IIIA–IIIB methods, these extra solutions correspond to higher frequency regions of the continuous dispersion relation and indicate that the discretisations are stable. Lastly, in Section 5.2, I have analysed the order in space of the explicit ODEs obtained in Chapter 3. An important consideration to be made when discretising a multi-Hamiltonian PDE in space by RK or PRK methods is that the active variables in the discretisation are the stage variables of the method, not the node variables (as is usually the case in the time integration of ODEs). The consequence of this is that the order of such a discretisation is restricted to the order of the stage variables in the RK or PRK method. For r-stage Gaussian RK methods, this order is $r + 1$, while for a Lobatto IIIA–IIIB method, the variables that Lobatto IIIA is applied to have order r. However, the ODEs that one obtains by discretising in space with an RK or PRK method contain approximations to the spatial derivatives, which, if they are not of the order of the stage variables, will restrict the order of the discretisation even further. For the explicit ODEs obtained in Chapter 3 from a Lobatto IIIA–IIIB discretisation in space of an appropriate multi-Hamiltonian PDE, I have shown that, for $r \leq 10$, the order of the approximation to the second derivatives (which are the only spatial derivatives that appear) is $r - 1$. I have conjectured that for higher values of r, the order of this approximation is also $r - 1$. Furthermore, I have shown that, for $r = 3$, there is a modification to the initial conditions that increases the order of the approximation to the second derivative. For higher values of r, similar modifications to the initial conditions that increase the order of this approximation are expected to exist.

6.2 Open questions

As with all research, the discovery of answers to old questions invariably leads to the discovery of new questions. Accordingly, from the research carried out during the tenure of my PhD and in the writing of this thesis comes a collection of new questions, a selection of which, I give below.

- As mentioned above, one of the themes that underlies much of this thesis is whether or not a multisymplectic discretisation of a multi-Hamiltonian PDE is well defined. In Chapter 2, it was shown that RK discretisations do not form well-defined multisymplectic integrators in general. In contrast, in Chapter 3, I have demonstrated that for a class of multi-Hamiltonian PDEs, discretisation in space by a Lobatto IIIA–IIIB method leads to explicit ODEs that do give a well-defined multisymplectic integrator (the integrator is explicit if the time discretisation is explicit). The properties of the Lobatto IIIA–IIIB methods that were required to demonstrate this are given in Eqs. (3.20), (3.21) and (3.22). This leads to the following questions:

  (i) What other symplectic PRK discretisations satisfy these three properties and how do the ODEs that they give compare with the ODEs from Lobatto IIIA–IIIB discretisations?

  (ii) What other PDEs can be written in multi-Hamiltonian form such that they satisfy the conditions of Theorem 3.4.1?

  (iii) For PRK discretisations that do not satisfy these three properties, can explicit ODEs in time be obtained by some other construction algorithm?

- A further requirement for a multisymplectic discretisation of a multi-Hamiltonian PDE to be well defined, which has been overlooked until now, is that the continuous PDE must be well posed to begin with. Therefore, it is necessary to ask, "What are the conditions on $\boldsymbol{K}$, $\boldsymbol{L}$ and $S(\boldsymbol{z})$ such that a multi-Hamiltonian PDE (1.6) (or its higher dimensional analogue, $\sum_{\alpha=1}^{N} \boldsymbol{K}^{\alpha}\partial_{x_\alpha}\boldsymbol{z} = \nabla_{\boldsymbol{z}} S(\boldsymbol{z})$) is well posed?"

- In Section 3.4.1, several examples of multi-Hamiltonian PDEs that do not satisfy the conditions of Theorem 3.4.1 are given. For these multi-Hamiltonian PDEs, multisymplectic box schemes may still be constructed in terms of the original variables in the PDE (see, for example, Eq. (3.70)). However, in these integrators, it is not possible to determine the value of all the variables that make up the multisymplectic conservation law. It is not clear what the implications of this are for the multisymplectic conservation law.

- The multisymplectic integrator formed for the NLS equation by a 2-stage Lobatto IIIA–IIIB discretisation in space and a linear–nonlinear splitting in time is shown to satisfy a non-local multisymplectic conservation law, since the numerical solution is non-local. However, for continuous multi-Hamiltonian PDEs, the conservation of multisymplecticity is a property that is local. This raises the question, "What role does locality play in the definition a multisymplectic conservation law?"

- The multi-Hamiltonian form of the BBM and Padé equations (given in Section 3.4.1) are new. Currently, there is no algorithm for determining the multi-Hamiltonian form of a PDE, or even whether such a multi-Hamiltonian form exists for a given

PDE. Thus, it is natural to ask, "What are the conditions on a PDE such that it may be written in multi-Hamiltonian form?"

- The order of the ODEs obtained in Chapter 3 is given by the minimum of the order of the stage variables and the order of the approximation to the second derivative. Conjecture 5.2.1 states that for an r-stage Lobatto IIIA–IIIB discretisation in space, this order is $r-1$. A proof of this conjecture would require showing that

$$\sum_{k=2}^{r-1}(\boldsymbol{C}^{-1})_{j-1,k-1}\big((1-c_k)(-c_j)^n-(c_k-c_j)^n+c_k(1-c_j)^n\big)=\begin{cases}1, & \text{for } n=2,\\ 0, & \text{for } 2<n<r+1,\end{cases} \tag{6.1}$$

for $j=2,\ldots,r-1$ (from Eq. (3.92)) and similarly for $j=1$ (from Eq. (3.93)). (For $n=0$ and $n=1$, the left-hand side of Eq. (6.1) is identically zero.)

- It was shown, in Section 5.2.1, that the order of the ODEs obtained from a 3-stage Lobatto IIIA–IIIB discretisation in space may be increased by an appropriate modification to the initial conditions. The exact form of this modification for Lobatto IIIA–IIIB methods with a different number of stages has not yet been determined. Furthermore, this leads to the question, "Does there exist a different modification to the initial conditions that would increase the order of the method even further?"

Appendix A

Proofs of various lemmas and theorems

A.1 Proof of Lemma 2.0.1

Proof. Since the Gaussian and Lobatto quadrature points are symmetric about $\frac{1}{2}$, they satisfy $c_i = 1 - c_{r+1-i}$. Now, $B(r)$ gives the relation

$$\sum_{i=1}^{r} b_i c_i^{k-1} = \frac{1}{k}, \qquad \text{for } k = 1, \ldots, r, \tag{A.1}$$

which is an expression, in terms of the standard polynomial basis, that the relation

$$\sum_{i=1}^{r} b_i \phi(c_i) = \int_0^1 \phi(x) \mathrm{d}x \tag{A.2}$$

holds for all polynomials $\phi(x)$ of degree less than r.

Let $\phi(x - \frac{1}{2})$ be any odd polynomial (i.e., $\phi(1-x) = -\phi(x)$) of degree less than r. Then,

$$\int_0^1 \phi(x) \mathrm{d}x = 0 \tag{A.3}$$

and Eq. (A.2) becomes

$$\sum_{i=1}^{\lfloor r/2 \rfloor} (b_i - b_{r+1-i}) \phi(c_i) = 0. \tag{A.4}$$

Therefore, we can conclude that $b_i = b_{r+1-i}$. □

A.2 Proof of Lemma 2.0.2

Unless otherwise noted, the indices in summations and products range from 1 to r.

Proof. First, Eq. (2.6) is rewritten as

$$\sum_j \boldsymbol{A}_{ij} \prod_{k \neq j} (c_j - c_k) = (-1)^{r-1} \prod_j c_j, \qquad \text{for } i = 1, \dots, r. \tag{A.5}$$

Then, expanding the product on the left gives

$$\begin{aligned} \sum_j \boldsymbol{A}_{ij} \prod_{k \neq j} (c_j - c_k) = \sum_j \boldsymbol{A}_{ij} \Big(c_j^{r-1} - c_j^{r-2} \sum_{k \neq j} c_k + \cdots \\ \cdots + (-1)^{l-1} c_j^{r-l} \sum_{\substack{k_1, \dots, k_{l-1}, \\ k_m \neq j \forall m, \\ k_{l-1} > \cdots > k_1}} \prod_{m=1}^{l-1} c_{k_m} + \cdots + \prod_{k \neq j} c_k \Big). \end{aligned} \tag{A.6}$$

Now,

$$\begin{aligned} (-1)^{l-1} c_j^{r-l} \sum_{\substack{k_1, \dots, k_{l-1}, \\ k_m \neq j \forall m, \\ k_{l-1} > \cdots > k_1}} \prod_{m=1}^{l-1} c_{k_m} = (-1)^{l-1} c_j^{r-l} \sum_{\substack{k_1, \dots, k_{l-1}, \\ k_{l-1} > \cdots > k_1}} \prod_{m=1}^{l-1} c_{k_m} \\ + (-1)^{l-2} c_j^{r-(l-1)} \sum_{\substack{k_1, \dots, k_{l-2}, \\ k_{l-2} > \cdots > k_1}} \prod_{m=1}^{l-2} c_{k_m} + \cdots - c_j^{r-2} \sum_{k_1} c_{k_1} + c_j^{r-1}. \end{aligned} \tag{A.7}$$

So,

$$\begin{aligned} \sum_j \boldsymbol{A}_{ij} \prod_{k \neq j} (c_j - c_k) = \sum_j \boldsymbol{A}_{ij} \Big(r c_j^{r-1} - (r-1) c_j^{r-2} \sum_k c_k + \cdots \\ \cdots + (-1)^{r-1} \sum_{\substack{k_1, \dots, k_{r-1}, \\ k_{r-1} > \cdots > k_1}} \prod_{m=1}^{r-1} c_{k_m} \Big). \end{aligned} \tag{A.8}$$

But GRK discretisations satisfy the simplifying assumption

$$\sum_i \boldsymbol{A}_{ij} c_j^{k-1} = \frac{1}{k} c_i^k, \qquad \text{for all } 1 \leq k \leq r. \tag{A.9}$$

Thus,

$$\begin{aligned} \sum_j \boldsymbol{A}_{ij} \prod_{k \neq j} (c_j - c_k) &= c_i^r - c_i^{r-1} \sum_k c_k + \cdots + (-1)^{r-1} c_i \sum_{\substack{k_1, \dots, k_{r-1}, \\ k_{r-1} > \cdots > k_1}} \prod_{m=1}^{r-1} c_{k_m} \\ &= (-1)^{r-1} \prod_k c_k. \end{aligned} \tag{A.10}$$

Therefore, if $\boldsymbol{A}$ is invertible, then Eq. (2.6) holds. $\square$

A.3 Proof of Lemma 3.4.2

Proof. From $D(\xi)$ with $k=1$ we have

$$\sum_{i=1}^{r} b_i A_{i,j}^{(2)} = b_j(1-c_j). \tag{A.11}$$

Since the Lobatto quadrature points satisfy $c_i = 1 - c_{r+1-i}$ and $b_j = b_{r+1-j}$ by Lemma 2.0.1, we have

$$\begin{aligned}\sum_{i=1}^{r} b_i A_{i,r+1-j}^{(2)} &= b_{r+1-j}(1-c_{r+1-j})\\ &= b_j c_j.\end{aligned} \tag{A.12}$$

Combining Eqs. (A.11) and (A.12) gives Eq. (3.97). □

A.4 Proof of Theorem 4.2.1

Proof. Eq. (3.14) for the NLS equation discretised in space by the 2-stage Lobatto IIIA–IIIB PRK discretisation with the partitioning $\boldsymbol{z}^{(1)} = \{p,q\}$ and $\boldsymbol{z}^{(2)} = \{v,w\}$ is given by

$$\Delta x \sum_{j=1}^{2} b_j \partial_t \omega_{i,j} + \kappa_{i+1} - \kappa_i = 0, \tag{A.13}$$

where $\omega_{i,j} = \mathrm{d}P_{i,j} \wedge \mathrm{d}Q_{i,j}$ and $\kappa_i = \mathrm{d}v_i \wedge \mathrm{d}p_i + \mathrm{d}w_i \wedge \mathrm{d}q_i$.

The first term in Eq. (A.13) as

$$\begin{aligned}\Delta x \sum_{j=1}^{2} b_j \partial_t \omega_{i,j} &= \frac{\Delta x}{2}\sum_{j=1}^{2} \partial_t(\mathrm{d}P_{i,j}\wedge \mathrm{d}Q_{i,j})\\ &= \frac{\Delta x}{2}\left(\partial_t(\mathrm{d}P_{i,1}\wedge \mathrm{d}Q_{i,1}) + \partial_t(\mathrm{d}P_{i,2}\wedge \mathrm{d}Q_{i,2})\right)\\ &= \frac{\Delta x}{2}\left(\partial_t(\mathrm{d}p_i \wedge \mathrm{d}q_i) + \partial_t(\mathrm{d}p_{i+1}\wedge \mathrm{d}q_{i+1})\right).\end{aligned} \tag{A.14}$$

Thus, we can write Eq. (A.13) as

$$\begin{aligned}&\frac{1}{2}\left(\partial_t(\mathrm{d}p_i \wedge \mathrm{d}q_i) + \partial_t(\mathrm{d}p_{i+1}\wedge \mathrm{d}q_{i+1})\right)\\ &\qquad + \frac{1}{\Delta x}\left(\mathrm{d}v_{i+1}\wedge \mathrm{d}p_{i+1} + \mathrm{d}w_{i+1}\wedge \mathrm{d}q_{i+1} - \mathrm{d}v_i \wedge \mathrm{d}p_i - \mathrm{d}w_i\wedge \mathrm{d}q_i\right) = 0.\end{aligned} \tag{A.15}$$

Now,

$$\mathrm{d}V_{i,1} = \mathrm{d}V_{i,2} = \mathrm{d}v_i + \frac{\Delta x}{2}(\partial_t \mathrm{d}Q_{i,1} - (6(P_{i,1})^2 + 2(Q_{i,1})^2)\mathrm{d}P_{i,1} - 4P_{i,1}Q_{i,1}\mathrm{d}Q_{i,1}) \tag{A.16}$$

and

$$\mathrm{d}p_{i+1} = \mathrm{d}p_i + \frac{\Delta x}{2}(\mathrm{d}V_{i,1} + \mathrm{d}V_{i,2}), \tag{A.17}$$

so

$$\begin{aligned}
\mathrm{d}v_{i+1} &= \mathrm{d}v_i + \frac{\Delta x}{2}\sum_j(\partial_t \mathrm{d}Q_{i,j} - (6(P_{i,j})^2 + 2(Q_{i,j})^2)\mathrm{d}P_{i,j} - 4P_{i,j}Q_{i,j}\mathrm{d}Q_{i,j}) \\
&= \mathrm{d}V_{i,1} + \frac{\Delta x}{2}(\partial_t \mathrm{d}q_{i+1} - (6(p_{i+1})^2 + 2(q_{i+1})^2)\mathrm{d}p_{i+1} - 4p_{i+1}q_{i+1}\mathrm{d}q_{i+1}) \qquad (A.18) \\
&= \frac{1}{\Delta x}(\mathrm{d}p_{i+1} - \mathrm{d}p_i) + \frac{\Delta x}{2}(\partial_t \mathrm{d}q_{i+1} - (6(p_{i+1})^2 + 2(q_{i+1})^2)\mathrm{d}p_{i+1} - 4p_{i+1}q_{i+1}\mathrm{d}q_{i+1})
\end{aligned}$$

and similarly for $\mathrm{d}w_{i+1}$.

Substituting for $\mathrm{d}v_{i+1}$, $\mathrm{d}v_i$, $\mathrm{d}w_{i+1}$ and $\mathrm{d}w_i$ in Eq. (A.15) gives

$$\begin{aligned}
&\frac{1}{2}\big(\partial_t(\mathrm{d}p_i \wedge \mathrm{d}q_i) + \partial_t(\mathrm{d}p_{i+1} \wedge \mathrm{d}q_{i+1})\big) \\
&\quad + \frac{1}{\Delta x}\Big(\big(\frac{1}{\Delta x}(\mathrm{d}p_{i+1} - \mathrm{d}p_i) \\
&\qquad + \frac{\Delta x}{2}(\partial_t \mathrm{d}q_{i+1} - (6(p_{i+1})^2 + 2(q_{i+1})^2)\mathrm{d}p_{i+1} - 4p_{i+1}q_{i+1}\mathrm{d}q_{i+1})\big) \wedge \mathrm{d}p_{i+1} \\
&\quad + \big(\frac{1}{\Delta x}(\mathrm{d}q_{i+1} - \mathrm{d}q_i) \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (A.19) \\
&\qquad + \frac{\Delta x}{2}(-\partial_t \mathrm{d}p_{i+1} - 4p_{i+1}q_{i+1}\mathrm{d}p_{i+1} - (2(p_{i+1})^2 + 6(q_{i+1})^2)\mathrm{d}q_{i+1})\big) \wedge \mathrm{d}q_{i+1} \\
&\quad - \big(\frac{1}{\Delta x}(\mathrm{d}p_i - \mathrm{d}p_{i-1}) + \frac{\Delta x}{2}(\partial_t \mathrm{d}q_i - (6(p_i)^2 + 2(q_i)^2)\mathrm{d}p_i - 4p_iq_i\mathrm{d}q_i)\big) \wedge \mathrm{d}p_i \\
&\quad - \big(\frac{1}{\Delta x}(\mathrm{d}q_i - \mathrm{d}q_{i-1}) + \frac{\Delta x}{2}(-\partial_t \mathrm{d}p_i - 4p_iq_i\mathrm{d}p_i - (2(p_i)^2 + 6(q_i)^2)\mathrm{d}q_i)\big) \wedge \mathrm{d}q_i\Big) = 0.
\end{aligned}$$

Since $\mathrm{d}p_i \wedge \mathrm{d}p_i = 0$ and $\mathrm{d}q_i \wedge \mathrm{d}p_i = -\mathrm{d}p_i \wedge \mathrm{d}q_i$, this simplifies to

$$\begin{aligned}
&\frac{1}{2}\big(\partial_t(\mathrm{d}p_i \wedge \mathrm{d}q_i) + \partial_t(\mathrm{d}p_{i+1} \wedge \mathrm{d}q_{i+1})\big) \\
&\qquad + \frac{1}{(\Delta x)^2}\left(\mathrm{d}p_{i+1} \wedge \mathrm{d}p_i + \mathrm{d}q_{i+1} \wedge \mathrm{d}q_i - \mathrm{d}p_i \wedge \mathrm{d}p_{i-1} - \mathrm{d}q_i \wedge \mathrm{d}q_{i-1}\right) \qquad (A.20) \\
&\qquad + \frac{1}{2}\left(-\mathrm{d}p_{i+1} \wedge \partial_t \mathrm{d}q_{i+1} - \partial_t \mathrm{d}p_{i+1} \wedge \mathrm{d}q_{i+1} + \mathrm{d}p_i \wedge \partial_t \mathrm{d}q_i + \partial_t \mathrm{d}p_i \wedge \mathrm{d}q_i\right) = 0,
\end{aligned}$$

which further simplifies to Eq. (4.10). □

A.5 Proof of Theorem 4.3.1

Proof. Eq. (3.7) for the NLS equation discretised in space and time by the 2-stage Lobatto IIIA–IIIB PRK discretisation with partitioning $\{(p,q),(v,w)\}$ in space and $\{(p,v),(q,w)\}$ in time is given by

$$\frac{1}{\Delta t}\sum_j b_j(\omega_j^{n+1}-\omega_j^n)+\frac{1}{\Delta x}\sum_m B_m(\kappa_{i+1}^m-\kappa_i^m)=0, \tag{A.21}$$

where

$$\begin{aligned}\omega_j^n &= \mathrm{d}P_{i,j}^n\wedge \mathrm{d}Q_{i,j}^n,\\ \kappa_i^m &= \mathrm{d}V_i^{n,m}\wedge \mathrm{d}P_i^{n,m}+\mathrm{d}W_i^{n,m}\wedge \mathrm{d}Q_i^{n,m}.\end{aligned} \tag{A.22}$$

Expanding the components of Eq. (A.21) gives

$$\begin{aligned}&\frac{1}{\Delta t}\sum_j b_j(\omega_j^{n+1}-\omega_j^n)=\\ &\qquad\frac{1}{2\Delta t}\Big(\mathrm{d}P_{i,1}^{n+1}\wedge \mathrm{d}Q_{i,1}^{n+1}+\mathrm{d}P_{i,2}^{n+1}\wedge \mathrm{d}Q_{i,2}^{n+1}-\mathrm{d}P_{i,1}^{n}\wedge \mathrm{d}Q_{i,1}^{n}-\mathrm{d}P_{i,2}^{n}\wedge \mathrm{d}Q_{i,2}^{n}\Big)\end{aligned} \tag{A.23}$$

and

$$\begin{aligned}&\frac{1}{\Delta x}\sum_m B_m(\kappa_{i+1}^m-\kappa_i^m)=\\ &\quad\frac{1}{2\Delta x}\left(\mathrm{d}V_{i+1}^{n,1}\wedge \mathrm{d}P_{i+1}^{n,1}+\mathrm{d}W_{i+1}^{n,1}\wedge \mathrm{d}Q_{i+1}^{n,1}-\mathrm{d}V_{i}^{n,1}\wedge \mathrm{d}P_{i}^{n,1}+\mathrm{d}W_{i}^{n,1}\wedge \mathrm{d}Q_{i}^{n,1}\right)\\ &\quad+\frac{1}{2\Delta x}\left(\mathrm{d}V_{i+1}^{n,2}\wedge \mathrm{d}P_{i+1}^{n,2}+\mathrm{d}W_{i+1}^{n,2}\wedge \mathrm{d}Q_{i+1}^{n,2}-\mathrm{d}V_{i}^{n,2}\wedge \mathrm{d}P_{i}^{n,2}+\mathrm{d}W_{i}^{n,2}\wedge \mathrm{d}Q_{i}^{n,2}\right).\end{aligned} \tag{A.24}$$

Now, since

$$\begin{aligned}\mathrm{d}V_{i,1}^{n,m}&=\mathrm{d}V_{i,2}^{n,m}=:\mathrm{d}V_{i,\frac12}^{n,m},\\ \mathrm{d}V_{i+1}^{n,m}&=\mathrm{d}V_{i,\frac12}^{n,m}+\frac{\Delta x}{2}\partial_x \mathrm{d}V_{i,2}^{n,m},\\ \mathrm{d}P_{i+1}^{n,m}&=\mathrm{d}P_{i}^{n,m}+\Delta x\,\mathrm{d}V_{i,\frac12}^{n,m},\end{aligned} \tag{A.25}$$

we can write

$$\begin{aligned}\mathrm{d}V_{i+1}^{n,m}&=\frac{1}{\Delta x}(\mathrm{d}P_{i+1}^{n,m}-\mathrm{d}P_{i}^{n,m})\\ &\quad+\frac{\Delta x}{2}(\partial_t \mathrm{d}Q_{i,2}^{n,m}-(6(P_{i,2}^{n,m})^2+2(Q_{i,2}^{n,m})^2)\mathrm{d}P_{i,2}^{n,m}-4P_{i,2}^{n,m}Q_{i,2}^{n,m}\mathrm{d}Q_{i,2}^{n,m}).\end{aligned} \tag{A.26}$$

Similarly, we can write

$$\begin{aligned}\mathrm{d}W_{i+1}^{n,m}&=\frac{1}{\Delta x}(\mathrm{d}Q_{i+1}^{n,m}-\mathrm{d}Q_{i}^{n,m})\\ &\quad+\frac{\Delta x}{2}(-\partial_t \mathrm{d}P_{i,2}^{n,m}-4P_{i,2}^{n,m}Q_{i,2}^{n,m}\mathrm{d}P_{i,2}^{n,m}-(2(P_{i,2}^{n,m})^2+6(Q_{i,2}^{n,m})^2)\mathrm{d}Q_{i,2}^{n,m}).\end{aligned} \tag{A.27}$$

Noting that

$$\begin{aligned} \mathrm{d}Q_i^{n,1} &= \mathrm{d}Q_i^{n,2} =: \mathrm{d}Q_i^{n,\frac{1}{2}}, \\ \mathrm{d}Q_{i,2}^{n,\frac{1}{2}} &= \mathrm{d}Q_{i+1}^{n,\frac{1}{2}}, \\ \mathrm{d}P_{i,2}^{n,m} &= \mathrm{d}P_{i+1}^{n,m}, \end{aligned} \tag{A.28}$$

we find that Eq. (A.24) becomes

$$\begin{aligned} \frac{1}{\Delta x}\sum_m B_m(\kappa_{i+1}^m - \kappa_i^m) = \frac{1}{2\Delta x^2}\Big(&\mathrm{d}P_{i+1}^{n,1} \wedge \mathrm{d}P_i^{n,1} + \mathrm{d}P_{i+1}^{n,2} \wedge \mathrm{d}P_i^{n,2} - \mathrm{d}P_i^{n,1} \wedge \mathrm{d}P_{i-1}^{n,1} \\ &- \mathrm{d}P_i^{n,2} \wedge \mathrm{d}P_{i-1}^{n,2} + 2(\mathrm{d}Q_{i+1}^{n,\frac{1}{2}} \wedge \mathrm{d}Q_i^{n,\frac{1}{2}} - \mathrm{d}Q_i^{n,\frac{1}{2}} \wedge \mathrm{d}Q_{i-1}^{n,\frac{1}{2}}) \\ &+ \frac{1}{4}\Big(\partial_t \mathrm{d}Q_{i+1}^{n,1} \wedge \mathrm{d}P_{i+1}^{n,1} + \partial_t \mathrm{d}Q_{i+1}^{n,2} \wedge \mathrm{d}P_{i+1}^{n,2} \\ &- \partial_t \mathrm{d}Q_i^{n,1} \wedge \mathrm{d}P_i^{n,1} - \partial_t \mathrm{d}Q_i^{n,2} \wedge \mathrm{d}P_i^{n,2} \\ &- (\partial_t \mathrm{d}P_{i+1}^{n,1} + \partial_t \mathrm{d}P_{i+1}^{n,2}) \wedge \mathrm{d}Q_{i+1}^{n,\frac{1}{2}} + (\partial_t \mathrm{d}P_i^{n,1} + \partial_t \mathrm{d}P_i^{n,2}) \wedge \mathrm{d}Q_i^{n,\frac{1}{2}}\Big)\Big), \end{aligned} \tag{A.29}$$

after cancelling terms of the form $\mathrm{d}Z \wedge \mathrm{d}Z$ and $\mathrm{d}P \wedge \mathrm{d}Q + \mathrm{d}Q \wedge \mathrm{d}P$.

Furthermore,

$$\begin{aligned} \partial_t \mathrm{d}Q_i^{n,1} &= \frac{2}{\Delta t}(\mathrm{d}Q_i^{n,\frac{1}{2}} - \mathrm{d}Q_i^n), \\ \partial_t \mathrm{d}Q_i^{n,2} &= \frac{2}{\Delta t}(\mathrm{d}Q_i^{n+1} - \mathrm{d}Q_i^{n,\frac{1}{2}}), \\ (\partial_t \mathrm{d}P_i^{n,1} + \partial_t \mathrm{d}P_i^{n,2}) &= \frac{2}{\Delta t}(\mathrm{d}P_i^{n+1} - \mathrm{d}P_i^n), \end{aligned} \tag{A.30}$$

so the last part of Eq. (A.29) becomes

$$\begin{aligned} &\frac{1}{4}\Big(\partial_t \mathrm{d}Q_{i+1}^{n,1} \wedge \mathrm{d}P_{i+1}^{n,1} + \partial_t \mathrm{d}Q_{i+1}^{n,2} \wedge \mathrm{d}P_{i+1}^{n,2} - \partial_t \mathrm{d}Q_i^{n,1} \wedge \mathrm{d}P_i^{n,1} - \partial_t \mathrm{d}Q_i^{n,2} \wedge \mathrm{d}P_i^{n,2} \\ &\qquad - (\partial_t \mathrm{d}P_{i+1}^{n,1} + \partial_t \mathrm{d}P_{i+1}^{n,2}) \wedge \mathrm{d}Q_{i+1}^{n,\frac{1}{2}} + (\partial_t \mathrm{d}P_i^{n,1} + \partial_t \mathrm{d}P_i^{n,2}) \wedge \mathrm{d}Q_i^{n,\frac{1}{2}}\Big) \\ &= \frac{1}{2\Delta t}\Big((\mathrm{d}Q_{i+1}^{n,\frac{1}{2}} - \mathrm{d}Q_{i+1}^n) \wedge \mathrm{d}P_{i+1}^n + (\mathrm{d}Q_{i+1}^{n+1} - \mathrm{d}Q_{i+1}^{n,\frac{1}{2}}) \wedge \mathrm{d}P_{i+1}^{n+1} \\ &\qquad - (\mathrm{d}Q_i^{n,\frac{1}{2}} - \mathrm{d}Q_i^n) \wedge \mathrm{d}P_i^n - (\mathrm{d}Q_i^{n+1} - \mathrm{d}Q_i^{n,\frac{1}{2}}) \wedge \mathrm{d}P_i^{n+1} \\ &\qquad - (\mathrm{d}P_{i+1}^{n+1} - \mathrm{d}P_{i+1}^n) \wedge \mathrm{d}Q_{i+1}^{n,\frac{1}{2}} + (\mathrm{d}P_i^{n+1} - \mathrm{d}P_i^n) \wedge \mathrm{d}Q_i^{n,\frac{1}{2}}\Big) \\ &= \frac{1}{2\Delta t}\Big(\mathrm{d}P_{i+1}^n \wedge \mathrm{d}Q_{i+1}^n - \mathrm{d}P_{i+1}^{n+1} \wedge \mathrm{d}Q_{i+1}^{n+1} - \mathrm{d}P_i^n \wedge \mathrm{d}Q_i^n + \mathrm{d}P_i^{n+1} \wedge \mathrm{d}Q_i^{n+1}\Big), \end{aligned} \tag{A.31}$$

which gives

$$\begin{aligned} \frac{1}{\Delta x}\sum_m B_m(\kappa_{i+1}^m - \kappa_i^m) = \frac{1}{2\Delta x^2}\Big(&\mathrm{d}P_{i+1}^n \wedge \mathrm{d}P_i^n + \mathrm{d}P_{i+1}^{n+1} \wedge \mathrm{d}P_i^{n+1} - \mathrm{d}P_i^n \wedge \mathrm{d}P_{i-1}^n \\ - \mathrm{d}P_i^{n+1} \wedge \mathrm{d}P_{i-1}^{n+1} &+ 2(\mathrm{d}Q_{i+1}^{n,\frac{1}{2}} \wedge \mathrm{d}Q_i^{n,\frac{1}{2}} - \mathrm{d}Q_i^{n,\frac{1}{2}} \wedge \mathrm{d}Q_{i-1}^{n,\frac{1}{2}})\Big) \\ + \frac{1}{2\Delta t}\Big(&\mathrm{d}P_{i+1}^n \wedge \mathrm{d}Q_{i+1}^n - \mathrm{d}P_{i+1}^{n+1} \wedge \mathrm{d}Q_{i+1}^{n+1} - \mathrm{d}P_i^n \wedge \mathrm{d}Q_i^n + \mathrm{d}P_i^{n+1} \wedge \mathrm{d}Q_i^{n+1}\Big). \end{aligned} \tag{A.32}$$

Combining Eqs. (A.23) and (A.32) we obtain

$$\begin{aligned}\frac{1}{\Delta t}\sum_j b_j(\omega_j^{n+1}-\omega_j^n)+\frac{1}{\Delta x}\sum_m B_m(\kappa_{i+1}^m-\kappa_i^m) &= \frac{1}{\Delta t}\left(\mathrm{d}P_i^{n+1}\wedge\mathrm{d}Q_i^{n+1}-\mathrm{d}P_i^n\wedge\mathrm{d}Q_i^n\right)\\ &+\frac{1}{2\Delta x^2}\Big(\mathrm{d}P_{i+1}^n\wedge\mathrm{d}P_i^n+\mathrm{d}P_{i+1}^{n+1}\wedge\mathrm{d}P_i^{n+1}-\mathrm{d}P_i^n\wedge\mathrm{d}P_{i-1}^n\\ &-\mathrm{d}P_i^{n+1}\wedge\mathrm{d}P_{i-1}^{n+1}+2(\mathrm{d}Q_{i+1}^{n,\frac12}\wedge\mathrm{d}Q_i^{n,\frac12}-\mathrm{d}Q_i^{n,\frac12}\wedge\mathrm{d}Q_{i-1}^{n,\frac12})\Big).\end{aligned}\tag{A.33}$$

Now, the variational equivalent of the last line of Eq. (4.12) is

$$\begin{aligned}\mathrm{d}Q_i^{n+1}=\mathrm{d}Q_i^{n,\frac12}+\frac{\Delta t}{2}\Big(&\frac{1}{\Delta x^2}(\mathrm{d}P_{i-1}^{n+1}-2\mathrm{d}P_i^{n+1}+\mathrm{d}P_{i+1}^{n+1})\\ &+(6(P_i^{n+1})^2+2(Q_i^{n,\frac12})^2)\mathrm{d}P_i^{n+1}+4P_i^{n+1}Q_i^{n,\frac12}\mathrm{d}Q_i^{n,\frac12}\Big).\end{aligned}\tag{A.34}$$

Substituting this into Eq. (A.33) and eliminating terms of the form $\mathrm{d}Z\wedge\mathrm{d}Z$ and $\mathrm{d}P\wedge\mathrm{d}Q+\mathrm{d}Q\wedge\mathrm{d}P$ gives

$$\begin{aligned}&\frac{1}{\Delta t}\sum_j b_j(\omega_j^{n+1}-\omega_j^n)+\frac{1}{\Delta x}\sum_m B_m(\kappa_{i+1}^m-\kappa_i^m)=\\ &\quad(\frac{1}{\Delta t}+2P_i^{n+1}Q_i^{n,\frac12})\mathrm{d}P_i^{n+1}\wedge\mathrm{d}Q_i^{n,\frac12}-(\frac{1}{\Delta t}+2P_i^nQ_i^{n-1,\frac12})\mathrm{d}P_i^{n+1}\wedge\mathrm{d}Q_i^{n-1,\frac12}\\ &\quad+\frac{1}{\Delta x^2}\Big((\mathrm{d}P_{i+1}^n+\mathrm{d}P_{i-1}^n)\wedge\mathrm{d}P_i^n+(\mathrm{d}Q_{i+1}^{n,\frac12}+\mathrm{d}Q_{i-1}^{n,\frac12})\wedge\mathrm{d}Q_i^{n,\frac12}\Big).\end{aligned}\tag{A.35}$$

Recalling that $\mathrm{d}p_i^n=\mathrm{d}P_{i,1}^{n,1}=\mathrm{d}P_i^n$ and defining $\mathrm{d}q_i^{n+\frac12}=\mathrm{d}Q_i^{n,\frac12}$, we can finally write

$$\begin{aligned}&\frac{1}{\Delta t}\sum_j b_j(\omega_j^{n+1}-\omega_j^n)+\frac{1}{\Delta x}\sum_m B_m(\kappa_{i+1}^m-\kappa_i^m)=\\ &\quad(\frac{1}{\Delta t}+2p_i^{n+1}q_i^{n+\frac12})\mathrm{d}p_i^{n+1}\wedge\mathrm{d}q_i^{n+\frac12}-(\frac{1}{\Delta t}+2p_i^nq_i^{n-\frac12})\mathrm{d}p_i^n\wedge\mathrm{d}q_i^{n-\frac12}\\ &\quad+\frac{1}{\Delta x^2}\Big((\mathrm{d}p_{i+1}^n+\mathrm{d}p_{i-1}^n)\wedge\mathrm{d}p_i^n+(\mathrm{d}q_{i+1}^{n+\frac12}+\mathrm{d}q_{i-1}^{n+\frac12})\wedge\mathrm{d}q_i^{n+\frac12}\Big),\end{aligned}\tag{A.36}$$

thus proving Theorem 4.3.1. □

Appendix B

Theorems and lemmas of [33]

The following definitions, theorems, lemmas and their proofs are quoted directly from [33] (with only adjustments to the equation numbers for readability). They show the relationship between collocation, discontinuous collocation and RK methods. Moreover, they give the order of the collocation and discontinuous collocation methods and the order of the collocation polynomial. While these results have a long history predating [33], for compactness and consistency of notation, it is convenient to present them as they are presented in [33].

The system under consideration in this Appendix is the non-autonomous system of first order ODEs

$$\dot{y} = f(t, y) \qquad y(t_0) = y_0. \tag{B.1}$$

Definition 1.1. Let b_i, a_{ij} $(i, j = 1, \ldots, s)$ be real numbers and let $c_i = \sum_{j=1}^{s} a_{ij}$. An *s-stage Runge–Kutta method* is given by

$$\begin{aligned} k_i &= f\Big(t_0 + c_i h, y_0 + h \sum_{j=1}^{s} a_{ijk_j}\Big), \qquad i = 1, \ldots, s \\ y_1 &= y_0 + h \sum_{i=1}^{s} b_i k_i. \end{aligned} \tag{B.2}$$

Definition 1.3. Let $c_1, \ldots, c_s$ be distinct real numbers (usually $0 \leq c_i \leq 1$). The *collocation polynomial* $u(t)$ is a polynomial of degree s satisfying

$$\begin{aligned} &u(t_0) = y_0 \\ &\dot{u}(t_0 + c_i h) = f(t_0 + c_i h, u(t_0 + c_i h)), \qquad i = 1, \ldots, s, \end{aligned} \tag{B.3}$$

and the numerical solution of the *collocation method* is defined by $y_1 = u(t_0 + h)$.

Theorem 1.4 (Guillou & Soulé 1969, Wright 1970). [32, 78] *The collocation method of Definition 1.3 is equivalent to the s-stage Runge–Kutta method (B.2) with coefficients*

$$a_{ij} = \int_0^{c_i} \ell_j(\tau)d\tau, \qquad b_i = \int_0^1 \ell_i(\tau)d\tau, \tag{B.4}$$

where $\ell_i(\tau)$ is the Lagrange polynomial $\ell_i(\tau) = \prod_{l\neq i}(\tau - c_l)/(c_i - c_l)$.

Proof. Let $u(t)$ be the collocation polynomial and define

$$k_i := \dot{u}(t_0 + c_i h). \tag{B.5}$$

By the Lagrange interpolation formula we have $\dot{u}(t_0 + c_i h) = \sum_{j=1}^{s} k_j \cdot \ell_j(\tau)$, and by integration we get

$$u(t_0 + c_i h) = y_0 + h \sum_{j=1}^{s} k_j \int_0^{c_i} \ell_j(\tau)d\tau. \tag{B.6}$$

Inserted into (B.3) this gives the first formula of the Runge–Kutta equation (B.2). Integration from 0 to 1 yields the second one. □

The above proof can also be read in reverse order. This shows that a Runge–Kutta method with coefficients given by (B.4) can be interpreted as a collocation method. Since $\tau^{k-1} = \sum_{j=1}^{s} c_j^{k-1}\ell_j(\tau)$ for $k = 1, \dots, s$, the relations (B.4) are equivalent to the linear systems

$$\begin{aligned} C(q): &\quad \sum_{j=1}^{s} a_{ij}c_j^{k-1} = \frac{c_i^k}{k}, \qquad k = 1, \dots, q, \text{ all } i \\ B(p): &\quad \sum_{i=1}^{s} b_i c_i^{k-1} = \frac{1}{k}, \qquad k = 1, \dots, p, \end{aligned} \tag{B.7}$$

with $q = s$ and $p = s$.

(Note that $B(q)$ and $C(p)$ are precisely the relations in Eq. (2.5) and Eq. (3.18) for $s = r$ and $q = p = \xi$.)

Theorem 1.5 (Superconvergence). *If the condition $B(p)$ holds for some $p \geq s$, then the collocation method (Definition 1.3) has order p. This means that the collocation method has the same order as the underlying quadrature formula.*

Proof. We consider the collocation polynomial $u(t)$ as the solution of a perturbed differential equation

$$\dot{u} = f(t, u) + \delta(t) \tag{B.8}$$

with defect $\delta(t) := \dot{u}(t) - f(t, u(t))$. Subtracting (B.1) from (B.8) we get after linearization that

$$\dot{u}(t) - \dot{y}(t) = \frac{\partial f}{\partial y}(t, y(t))\,(u(t) - y(t)) + \delta(t) + r(t), \tag{B.9}$$

where, for $t_0 \leq t \leq t_0 + h$, the remainder $r(t)$ is $\mathcal{O}(\|u(t) - y(t)\|^2) = \mathcal{O}(h^{2s+2})$ by Lemma 1.6 below. The variation of constants formula (see Hairer, Nørsett & Wanner (1993), p. 66 [35]) then yields

$$y_1 - y(t_0 + h) = u(t_0 + h) - y(t_0 + h) = \int_{t_0}^{t_0+h} R(t_0 + h, s)\,(\delta(s) + r(s))\,ds, \qquad \text{(B.10)}$$

where $R(t, s)$ is the resolvent of the homogeneous part of the differential equation (B.9), i.e., the solution of the matrix differential equation $\partial R(t, s)/\partial t = A(t)R(t, s)$, $R(s, s) = I$, with $A(t) = \partial f/\partial y(t, y(t))$. The integral over $R(t_0 + h, s)r(s)$ gives a $\mathcal{O}(h^{2s+3})$ contribution. The main idea now is to apply the quadrature formula $(b_i, c_i)_{i=1}^s$ to the integral over $g(s) = R(t_0 + h, s)\delta(s)$; because the defect $\delta(s)$ vanishes at the collocation points $t_0 + c_i h$ for $i = 1, \ldots, s$, this gives zero as the numerical result. Thus, the integral is equal to the quadrature error, which is bounded by h^{p+1} times a bound of the pth derivative of the function $g(s)$. This derivative is bounded independently of h, because by Lemma 1.6 all derivatives of the collocation polynomial are bounded uniformly as $h \to 0$. Since, anyway, $p \leq 2s$, we get $y_1 - y(t_0 + h) = \mathcal{O}(h^{p+1})$ from (B.10). □

Lemma 1.6. *The collocation polynomial $u(t)$ is an approximation of order s to the exact solution of (B.1) on the whole interval, i.e.,*

$$\|u(t) - y(t)\| \leq C \cdot h^{s+1} \qquad \textit{for } t \in [t_0, t_0 + h] \qquad \text{(B.11)}$$

and for sufficiently small h.

Moreover, the derivatives of $u(t)$ satisfy for $t \in [t_0, t_0 + h]$

$$\|u^{(k)}(t - y^{(k)}(t))\| \leq C \cdot h^{s+1-k} \qquad \textit{for } k = 0, \ldots, s. \qquad \text{(B.12)}$$

Proof. The collocation polynomial satisfies

$$\dot{u}(t_0 + \tau h) = \sum_{i=1}^{s} f\left(t_0 + c_i h, u(t_0 + c_i h)\right) \ell_i(\tau), \qquad \text{(B.13)}$$

while the exact solution of (B.1) satisfies

$$\dot{y}(t_0 + \tau h) = \sum_{i=1}^{s} f\left(t_0 + c_i h, y(t_0 + c_i h)\right) \ell_i(\tau) + h^s E(\tau, h), \qquad \text{(B.14)}$$

where the interpolation error $E(\tau, h)$ is bounded by $\max_{t \in [t_0, t_0+h]} \|y^{(s+1)}(t)\|/s!$ and its derivatives satisfy

$$\|E^{(k-1)}(\tau,h)\| \leq \max_{t\in[t_0,t_0+h]} \frac{\|y^{(s+1)}(t)\|}{(s-k+1)!}. \tag{B.15}$$

This follows from the fact that, by Rolle's theorem, the differential polynomial $\sum_{i=1}^{s} f\left(t_0+c_ih, y(t_0+c_ih)\right)\ell_i^{k-1}(\tau)$ can be interpreted as the interpolation polynomial of $h^{k-1}y^{(k)}(t_0+\tau h)$ at $s-k+1$ points lying in $[t_0, t_0+h]$. Integrating the difference of the above two equations gives

$$y(t_0+\tau h)-u(t_0+\tau h) = h\sum_{i=1}^{s}\Delta f_i\int_0^{\tau}\ell_i(\sigma)d\sigma + h^{s+1}\int_0^{\tau}E(\sigma,h)d\sigma \tag{B.16}$$

with $\Delta f_i = f\left(t_0+c_ih, y(t_0+c_ih)\right) - f\left(t_0+c_ih, u(t_0+c_ih)\right)$. Using a Lipschitz condition for $f(t,y)$, this relation yields

$$\max_{t\in[t_0,t_0+h]}\|y(t)-u(t)\| \leq hCL\max_{t\in[t_0,t_0+h]}\|y(t)-u(t)\| + Const\cdot h^{s+1}, \tag{B.17}$$

implying the statement (B.11) for sufficiently small $h>0$.

The proof of the second statement follows from

$$h^k\left(y^{(k)}(t_0+\tau h)-u^{(k)}(t_0+\tau h)\right) = h\sum_{i=1}^{s}\Delta f_i\ell_i^{k-1}(\tau) + h^{s+1}E^{(k-1)}(\tau,h) \tag{B.18}$$

by using a Lipschitz condition for $f(t,y)$ and the estimate (B.11). □

Definition 1.7. Let $c_2,\ldots,c_{s-1}$ be distinct real numbers (usually $0\leq c_i\leq 1$), and let b_1, b_s be two arbitrary real numbers. The corresponding *discontinuous collocation method* is then defined via a polynomial of degree $s-2$ satisfying

$$\begin{aligned}
&u(t_0) = y_0 - hb_1\left(\dot{u}(t_0) - f(t_0,u(t_0))\right)\\
&\dot{u}(t_0+c_ih) = f(t_0|c_ih, u(t_0+c_ih)), \qquad i=2,\ldots,s-1,\\
&y_1 = u(t_1) - hb_s\left(\dot{u}(t_1)-f(t_1,u(t_1))\right).
\end{aligned} \tag{B.19}$$

and the numerical solution of the *collocation method* is defined by $y_1 = u(t_0+h)$.

Theorem 1.8. *The discontinuous collocation method of Definition 1.7 is equivalent to an s-stage Runge–Kutta method (B.2) with coefficients determined by* $c_1=0$, $c_s=1$, *and*

$$\begin{aligned}
&a_{i1}=b_1, \qquad a_{is}=0 \qquad \text{for } i=1,\ldots,s,\\
&C(s-2) \qquad \text{and} \qquad B(s-2),
\end{aligned} \tag{B.20}$$

with conditions $C(q)$ and $B(p)$ of Eq. (B.7).

Proof. As in the proof of Theorem 1.4 we put $k_i := \dot{u}(t_0+c_ih)$ (this time for $i = 2, \dots, s-1$), so that $\dot{u}(t_0 + \tau h) = \sum_{j=2}^{s-1} k_j \cdot \ell_j(\tau)$ by the Lagrange interpolation formula. Here, $\ell_j(\tau)$ corresponds to $c_2, \dots, c_{s-1}$ and is a polynomial of degree $s-3$. By integration and using the definition of $u(t_0)$ we get

$$\begin{aligned} u(t_0 + c_ih) &= u(t_0) + h \sum_{j=2}^{s-1} k_j \int_0^{c_i} \ell_j(\tau) d\tau \\ &= y_0 + hb_1k_1 + h \sum_{j=2}^{s-1} k_j \left(\int_0^{c_i} \ell_j(\tau) d\tau - b_1 \ell_j(0) \right) \end{aligned} \tag{B.21}$$

with $k_1 = f(y_0)$. Inserted into (B.19) this gives the first formula of the Runge–Kutta equation (B.2) with $a_{ij} = \int_0^{c_i} \ell_j(\tau)d\tau - b_1\ell_j(0)$. As for collocation methods, one checks that the a_{ij} are uniquely determined by the condition $C(s-2)$. The formula for y_1 is obtained similarly. $\square$

Theorem 1.9 (Superconvergence). *The discontinuous collocation method of Definition 1.7 has the same order as the underlying quadrature formula.*

Proof. We follow the same lines of Theorem 1.5. With the polynomial $u(t)$ of Definition 1.7, and with the defect

$$\delta(t) := \dot{u}(t) - f(t, u(t)) \tag{B.22}$$

we get (B.9) after linearization. The variation of constants formula then yields

$$u(t_0+h) - y(t_0+h) = R(t_0+h, t_0)(u(t_0) - y_0) + \int_{t_0}^{t_0+h} R(t_0+h, s) \left(\delta(s) + r(s) \right) ds, \tag{B.23}$$

which corresponds to (B.10) if $u(t_0) = y_0$. As a consequence of Lemma 1.10 below (with $k = 0$), the integral over $R(t_0 + h, s)r(s)$ gives a $\mathcal{O}(h^{2s-1})$ contribution. Since the defect $\delta(t_0 + c_ih)$ vanishes only for $i = 2, \dots, s-1$, an application of the quadrature formula to $R(t_0 + h, s)\delta(s)$ yields $hb_1R(t_0 + h, t_0)\delta(t_0) + hb_s\delta(t_0 + h)$ in addition to the quadrature error, which is $\mathcal{O}(h^{p+1})$. Collecting terms suitably, we obtain

$$u(t_1) - hb_s\delta(t_1) - y(t_1) = R(t_1, t_0) \left(u(t_0) + hb_1\delta(t_0) - y_0 \right) + \mathcal{O}(h^{p+1}) + \mathcal{O}(h^{2s-1}), \tag{B.24}$$

which, after using the definitions of $u(t_0)$ and $u(t_1)$, proves $y_1 - y(t_1) = \mathcal{O}(h^{p+1}) + \mathcal{O}(h^{2s-1})$. $\square$

Lemma 1.10. *The polynomial $u(t)$ of the discontinuous collocation method (B.19) satisfies for $t \in [t_0, t_0 + h]$ and for sufficiently small h*

$$\|u^{(k)}(t) - y^{(k)}(t)\| \le C \cdot h^{s-1-k} \qquad \textit{for } k = 0, \dots, s-2. \tag{B.25}$$

Proof. The proof is essentially the same as that for Lemma 1.6. In the formulas for $\dot{u}(t_0+\tau h)$ and $\dot{y}(t_0+\tau h)$, the sum has to be taken from $i=2$ to $i=s-1$. Moreover, all h^s become h^{s-2}. In (B.16) one has an additional term

$$y_0 - u(t_0) = hb_1 \left(\dot{u}(t_0) - f(t_0, u(t_0))\right), \tag{B.26}$$

which, however, is just an interpolation error of size $\mathcal{O}(h^{s-1})$ and can be included in $Const \cdot h^{s-1}$. □

Theorem 2.2. *The partitioned Runge–Kutta method composed of the s-stage Lobatto IIIA and the s-stage Lobatto IIIB method, is of order $2s-2$.*

Proof. Let $c_1 = 0$, $c_2, \ldots, c_{s-1}$, $c_s = 1$ and $b_1, \ldots, b_s$ be the nodes and weights of the Lobatto quadrature. The partitioned Runge–Kutta method based on the Lobatto IIIA–IIIB pair can be interpreted as the discontinuous collocation method

$$\begin{aligned}
&u(t_0) = y_0 \\
&v(t_0) = z_0 - hb_1\left(\dot{v}(t_0) - g(u(t_0), v(t_0))\right) \\
&\dot{u}(t_0 + c_i h) = f\left(u(t_0 + c_i h), v(t_0 + c_i h)\right), \qquad i = 1, \ldots, s \\
&\dot{v}(t_0 + c_i h) = g\left(u(t_0 + c_i h), v(t_0 + c_i h)\right), \qquad i = 2, \ldots, s-1 \\
&y_1 = u(t_1) \\
&z_1 = v(t_i) - hb_s\left(\dot{v}(t_1) - g(u(t_1), v(t_1))\right),
\end{aligned} \tag{B.27}$$

where $u(t)$ and $v(t)$ are polynomials of degree s and $s-2$, respectively. This is seen as in the proofs of Theorem 1.4 and Theorem 1.8. The superconvergence (order $2s-2$) is obtained with exactly the same proof as for Theorem 1.9, where the functions $u(t)$ and $y(t)$ have to be replaced with $(u(t), v(t))^T$ and $(y(t), z(t))^T$, etc. Instead of Lemma 1.10 we use the estimates (for $t \in [t_0, t_0 + h]$)

$$\begin{aligned}
&\|u^{(k)}(t) - y^{(k)}(t)\| \le c \cdot h^{s-k} \qquad \text{for } k = 0, \ldots, s, \\
&\|v^{(k)}(t) - z^{(k)}(t)\| \le c \cdot h^{s-1-k} \qquad \text{for } k = 0, \ldots, s-2,
\end{aligned} \tag{B.28}$$

which can be proved by following the lines of the proofs of Lemma 1.6 and Lemma 1.10. □

Bibliography

[1] L. Abia and J. M. Sanz-Serna. Partitioned Runge–Kutta methods for separable Hamiltonian systems. *Math. Comp.*, 60(202):617–634, 1993.

[2] U. M. Ascher and R. I. McLachlan. Multisymplectic box schemes and the Korteweg–de Vries equation. *Appl. Numer. Math*, 48:255–269, 2004.

[3] U. M. Ascher and R. I. McLachlan. On symplectic and multisymplectic schemes for the Korteweg–de Vries equation. *J. Sci. Comput.*, 25(1):83–104, 2005.

[4] T. B. Benjamin, J. L. Bona, and J. J. Mahony. Model equations for long waves in nonlinear dispersive systems. *Philos. Trans. Roy. Soc. London Ser. A*, 272(1220):47–78, 1972.

[5] A. I. Bobenko and Y. B. Suris. Discrete Lagrangian reduction, discrete Euler–Poincaré equations, and semidirect products. *Lett. Math. Phys.*, 49:79–93, 1999.

[6] A. I. Bobenko and Y. B. Suris. Discrete time Lagrangian mechanics on Lie groups, with and application to the Lagrange top. *Comm. Math. Phys.*, 204:147–188, 1999.

[7] T. J. Bridges. A geometric formulation of the conservation of wave action and its implications for signature and the classification of instabilities. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 453:1365–1395, 1997.

[8] T. J. Bridges. Multi-symplectic structures and wave propagation. *Math. Proc. Cambridge Philos. Soc.*, 121:147–190, 1997.

[9] T. J. Bridges and G. Derks. Unstable eigenvalues and the linearization about solitary waves and fronts with symmetry. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 455:2427–2469, 1999.

[10] T. J. Bridges, P. E. Hydon, and S. Reich. Vorticity and symplecticity in Lagrangian fluid dynamics. *J. Phys. A*, 38:1403–1418, 2005.

[11] T. J. Bridges and F. E. Laine-Pearson. Multisymplectic relative equilibria, multiphase wavetrains, and coupled NLS equations. *Stud. Appl. Math.*, 107:137–155, 2001.

[12] T. J. Bridges and F. E. Laine-Pearson. Nonlinear counterpropagating waves, multi-symplectic geometry, and the instability of standing waves. *SIAM J. Appl. Math.*, 64(6):2096–2120, 2004.

[13] T. J. Bridges and S. Reich. Multi-symplectic integrators: numerical schemes for Hamiltonian PDEs that conserve symplecticity. *Phys. Lett. A*, 284:184–193, 2001.

[14] T. J. Bridges and S. Reich. Multi-symplectic spectral discretizations for the Zakharov–Kuznetsov and shallow water equations. *Phys. D*, 152:491–504, 2001.

[15] T. J. Bridges and S. Reich. Numerical methods for Hamiltonian PDEs. *J. Phys. A*, 39:5287–5320, 2006.

[16] J. C. Butcher. Implicit Runge–Kutta processes. *Math. Comp.*, 18(85):50–64, 1964.

[17] J. B. Chen. Multisymplectic geometry, local conservation laws and Fourier pseudospectral discretization for the "good" Boussinesq equation. *Appl. Math. Comput.*, 161:55–67, 2005.

[18] J. B. Chen and M. Z. Qin. Multi-symplectic Fourier pseudospectral method for the nonlinear Schrödinger equation. *Electron. Trans. Numer. Anal.*, 12:193–204, 2001.

[19] F. H. Chipman. A–stable Runge–Kutta processes. *Nordisk Tidskr. Informationsbehandling (BIT)*, 11:384–388, 1971.

[20] T. De Donder. *Théorie Invariantive du Calcul de Variations.* Gauthier-Villars, Paris, 1935.

[21] P. Dedecker. Calcul des variations, formes différentielles et champs geodesiques. *Geometrie différentielle, Colloq. Intern. du CNRS LII, Strasbourg*, pages 17–34, 1953.

[22] P. Dedecker. On the generalization of symplectic geometry to multiple integrals in the calculus of variations. *Lecture Notes in Math.*, 570:395–456, 1977.

[23] S. Edvardsson, K. G. Karlsson, and M. Engholm. Accurate spin axes and solar system dynamics: Climatic variations for the earth and mars. *Astronom. and Astrophys.*, 384:689–701, 2002.

[24] B. L. Ehle. High order A-stable methods for the numerical solution of systems of D.E.'s. *Nordisk Tidskr. Informationsbehandling (BIT)*, 8:276–278, 1968.

[25] R. Fetecau and D. Levy. Approximate model equations for water waves. *Commun. Math. Sci.*, 3(2):159–170, 2005.

[26] J. Frank. Geometric space–time integration of ferromagnetic materials. *Appl. Numer. Math.*, 48:307–322, 2004.

[27] J. Frank, B. E. Moore, and S. Reich. Linear PDEs and numerical methods that preserve a multisymplectic conservation law. *SIAM J. Sci. Comput.*, 28(1):260–277, 2006.

[28] M. J. Gotay. *A multisymplectic framework for classical field theory and the calculus of variations I: Covariant Hamiltonian formalism*, pages 203–235. North Holland, Amsterdam, 1991.

[29] M. J. Gotay. A multisymplectic framework for classical field theory and the calculus of variations II: Space + time decomposition. *Differential Geom. Appl.*, 1:375–390, 1991.

[30] M. J. Gotay, J. Isenberg, R. Montgomery, J. E. Marsden, J. Śniatycki, and P. B. Yasskin. Momentum maps and classical relativistic fields. part I: Covariant field theory. *Arxiv preprint physics/9801019v2*, 2004.

[31] S. Gray and R. I. McLachlan. Optimal stability polynomials for splitting methods, with application to the time-dependent Schröder equation. *Appl. Numer. Anal.*, 25:275–286, 1997.

[32] A. Guillou and J. L. Soulé. La résolution numérique des problèmes différentiels aux conditions initiales par des méthodes de collocation. *Rev. Française Informat. Recherche Opérationnelle*, 3:17–44, 1969. (French).

[33] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration*. Springer–Verlag, 2002.

[34] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration*. Springer–Verlag, Berlin, second edition, 2006.

[35] E. Hairer, S. P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I. Nonstiff Problems*. Springer, Berlin, second edition, 1993.

[36] R. H. Hardin and F. D. Tappert. Applications of the split-step Fourier method to the numerical solution of nonlinear and variable coefficient wave equations. *SIAM Rev.*, 15:423, 1973.

[37] F. Hélein and J. Kouneiher. Covariant Hamiltonian formalism for the calculus of variations with several variables: Lepage–Dedecker versus de Donder–Weyl. *Adv. Theor. Math. Phys.*, 8(3):565–601, 2004.

[38] J. L. Hong, Y. Liu, and G. Sun. The multi-symplecticity of partitioned Runge–Kutta methods for Hamiltonian PDEs. *Math. Comp.*, 75:167–181, 2005.

[39] P. E. Hydon. Multisymplectic conservation laws for differential and differential-difference equations. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 461:1627–1637, 2005.

[40] P. E. Hydon. Vorticity and symplecticity in Lagrangian fluid dynamics. *J. Phys. A*, 38:1403–1418, 2005.

[41] A. L. Islas and C. M. Schober. Multisymplectic spectral methods for the Gross–Pitaevskii equation. In *Computational science - ICCS 2002, Part III (Amsterdam)*, volume 2331 of *Lecture Notes in Comput. Sci.*, pages 486–495, 2002.

[42] A. L. Islas and C. M. Schober. Multi-symplectic methods for generalized Schrödinger equations. *Future Generation Computer Systems*, 19:403–413, 2003.

[43] A. L. Islas and C. M. Schober. Multi-symplectic spectral methods for the sine–Gordon equation. In *Computational Science - ICCS 2003, Pt II*, volume 2658 of *Lecture Notes In Comput. Sci.*, pages 101–110, 2003.

[44] L. O. Jay. Symplectic partitioned Runge–Kutta methods for constrained Hamiltonian systems. *SIAM J. Numer. Anal.*, 33(1):368–387, 1996.

[45] C. Kane, J. E. Marsden, M. Ortiz, and M. West. Variational integrators and the Newmark algorithm for conservative and dissipative mechanical systems. *Internat. J. Numer. Methods in Engrg.*, 49:1295–1325, 2000.

[46] H. A. Kastrup. Canonical theories of Lagrangian dynamical systems in physics. *Phys. Rep.*, 101:1–167, 1983.

[47] J. Kijowski and W. M. Tulczyjew. *A symplectic framework for field theories*, volume 107 of *Lecture Notes in Physics*. Springer–Verlag, Berlin–New York, 1979.

[48] F. M. Lasagni. Canonical Runge–Kutta methods. *Z. Angew. Math. Phys.*, 39:952–953, 1988.

[49] B. Leimkuhler and S. Reich. *Simulating Hamiltonian Dynamics*. Cambridge University Press, Cambridge, 2004.

[50] A. Lew, J. E. Marsden, M. Ortiz, and M. West. Asynchronous variational integrators. *Arch. Ration. Mech. Anal.*, 167:85–146, 2003.

[51] T. Liu and M. Qin. Multisymplectic geometry and multisymplectic Preissman scheme for the KP equation. *J. Math. Phys.*, 43(8):4060–4077, 2002.

[52] J. E. Marsden, G. W. Patrick, and S. Shkoller. Multisymplectic geometry, variational integrators, and nonlinear PDEs. *Comm. Math. Phys.*, 199:351–395, 1998.

[53] J. E. Marsden, S. Pekarsky, and S. Shkoller. Symmetry reduction of discrete Lagrangian mechanics on Lie groups. *J. Geom. Phys.*, 36:139–150, 2000.

[54] J. E. Marsden, S. Pekarsky, S. Shkoller, and M. West. Variational methods, multisymplectic geometry and continuum mechanics. *J. Geom. Phys.*, 38:253–284, 2001.

[55] J. E. Marsden and T. S. Ratiu. *Introduction to Mechanics and Symmetry*. Springer–Verlag New York, Inc., 2nd edition, 1999.

[56] R. I. McLachlan, J. Frank, and A. Zanna. Variational integrators for the Landau–Lifshitz equation. Lecture presented by R. I. McLachlan at the *Geometric and Structure Preserving Algorithms for PDEs* workshop as part of the *Special Year in Geometric Integration*, Oslo, Norway, May 2003.

[57] R. I. McLachlan and G. R. W. Quispel. Splitting methods. *Acta Numer.*, 11:341–434, 2002.

[58] B. E. Moore. *A Modified Equations Approach for Multisymplectic Integration Methods*. PhD thesis, University of Surrey, 2003.

[59] B. E. Moore and S. Reich. Backward error analysis for multi-symplectic integration methods. *Numer. Math.*, 95:625–652, 2003.

[60] B. E. Moore and S. Reich. Multi-symplectic integration methods for Hamiltonian PDEs. *Future Generation Computer Systems*, 19(21):395–402, 2003.

[61] M. Oliver, M. West, and C. Wulff. Approximate momentum conservation for spatial semidiscretizations of semilinear wave equations. *Numer. Math.*, 97:493–535, 2004.

[62] C. Paufler and H. Römer. De Donder–Weyl equations and multisymplectic geometry. *Rep. Math. Phys.*, 49:325–334, 2002.

[63] S. Reich. Multi-symplectic Runge–Kutta collocation methods for Hamiltonian wave equations. *J. Comput. Phys.*, 157:473–499, 1999.

[64] S. Reich. Finite volume methods for multi-symplectic PDEs. *BIT*, 40:559–582, 2000.

[65] B. N. Ryland and R. I. McLachlan. On multisymplecticity of partitioned Runge–Kutta methods. *SIAM J. Sci. Comput.*, In press.

[66] B. N. Ryland, R. I. McLachlan, and J. Frank. On multisymplecticity of partitioned Runge–Kutta and splitting methods. *Int. J. Comput. Math.*, 84(6):847–869, 2007.

[67] J. M. Sanz-Serna. Runge–Kutta schemes for Hamiltonian systems. *BIT*, 28:877–883, 1988.

[68] Z. Shang. KAM theorem of symplectic algorithms for Hamiltonian systems. *Numer. Math.*, 83(3):477–496, 1999.

[69] H. Su, M. Qin, and R. Scherer. A multisymplectic geometry and a multisymplectic scheme for Maxwell's equations. *Int. J. Pure Appl. Math.*, 34(1):1–17, 2007.

[70] G. Sun. Symplectic partitioned Runge–Kutta methods. *J. Comput. Math.*, 11(4):365–372, 1993.

[71] Y. J. Sun and M. Z. Qin. Construction of multisymplectic schemes of any finite order for modified wave equations. *J. Math. Phys.*, 41(11):7854–7868, 2000.

[72] Y. B. Suris. Preservation of symplectic structure in the numerical solution of Hamiltonian systems. In S. S. Filippov, editor, *Numerical Solution of Ordinary Differential Equations*, pages 148–160. Akad. Nauk SSSR, Inst. Prikl. Mat., Moscow, 1988. (Russian).

[73] Y. B. Suris. Canonical transformations generated by methods of Runge–Kutta type for the numerical integration of the system $x'' = -\partial u/\partial x$. *Zh. Vychisl. Mat. i Mat. Fiz.*, 29:202–211, 1989. (Russian).

[74] J. von Rieth. The Hamilton–Jacobi theory of De Donder and Weyl applied to some relativistic field theories. *J. Math. Phys*, 25(4):1102–1115, 1984.

[75] J. M. Wendlandt and J. E. Marsden. Mechanical integrators derived from a discrete variational principle. *Phys. D*, 106:223–246, 1997.

[76] H. Weyl. Geodesic fields in the calculus of variation for multiple integrals. *Ann. of Math. (2)*, 36(3):607–629, 1935.

[77] H. Weyl. *The Classical Groups*. Princeton University Press, Princeton, N.J., 1939.

[78] K. Wright. Some relationships between implicit Runge–Kutta, collocation Lanczos τ methods, and their stability properties. *Nordisk Tidskr. Informationsbehandling (BIT)*, 10:217–227, 1970.

[79] H. Yoshida. Construction of higher order symplectic integrators. *Phys. Lett. A*, 150(5–7):262–268, 1990.

[80] P. F. Zhao and M. Z. Qin. Multisymplectic geometry and multisymplectic Preissmann scheme for the KdV equation. *J. Phys. A*, 33:3613–3626, 2000.