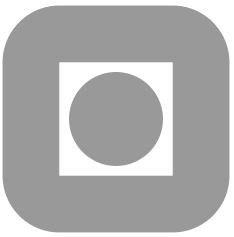
NORGES TEKNISK-NATURVITENSKAPELIGE
UNIVERSITET

# A general framework for deriving integral preserving numerical methods for PDEs

by

Morten Dahlby and Brynjulf Owren

PREPRINT
NUMERICS NO. 8/2010



NORWEGIAN UNIVERSITY OF
SCIENCE AND TECHNOLOGY
TRONDHEIM, NORWAY

This report has URL
http://www.math.ntnu.no/preprint/numerics/2010/N8-2010.pdf
Address: Department of Mathematical Sciences, Norwegian University of Science and Technology, N-7491 Trondheim, Norway.

# A general framework for deriving integral preserving numerical methods for PDEs

Morten Dahlby and Brynjulf Owren

September 17, 2010

A general procedure for constructing conservative numerical integrators for time dependent partial differential equations is presented. In particular, linearly implicit methods preserving a time averaged version of the invariant is developed for systems of partial differential equations with polynomial nonlinearities. The framework is rather general and allows for an arbitrary number of dependent and independent variables with derivatives of any order. It is proved formally that second order convergence is obtained. The procedure is applied to a test case and numerical experiments are provided.

## 1 Introduction

Schemes that conserve geometric structure have been shown to be useful when studying the long time behaviour of dynamical systems. Such schemes are sometimes called geometric or structure preserving integrators [16, 17]. In this paper we shall mostly be concerned with the conservation of first integrals.

Even if a presumption in this work is that the development of new and better integral preserving schemes is useful, we mention some situations where this presumption can be easily asserted. In the literature one finds several examples where stability of a numerical method is proved by directly using its conservative property, one example is the scheme developed for the cubic Schrödinger equation in [10]. Another application where the exact preservation of first integrals plays an important role is in the study of orbital stability of soliton solutions to certain Hamiltonian partial differential equations (PDEs) as discussed by Benjamin and coauthors [1, 2].

For ordinary differential equations (ODEs) it is common to devise relatively general frameworks for structure preservation. This is somewhat to the contrary of the usual practice with partial differential equations (PDEs) where each equation under consideration normally requires a dedicated scheme. But there exists certain fairly general methodologies that can be used for developing geometric schemes also for PDEs. For example, through space discretisation of a Hamiltonian PDE one may obtain a system of Hamiltonian ODEs to which a geometric integrator may be applied. Another approach is to formulate the PDE in multi-symplectic form, and then apply a scheme which preserves a discrete version of this form, see [3] for a review.

In this paper we consider methods for PDEs that are based on the discrete gradient method, which is one of the most popular methods for designing integral preserving schemes for ODEs. The discrete gradient method was perhaps first discussed by Gonzalez [15], but see also [16, 22]. For PDEs one may derive discrete gradients either for the abstract

Cauchy problem, where the solution at any time is considered as an element of some infinite dimensional space, or one may semidiscretise the equations in space and then derive the corresponding discrete gradient for the resulting ODE system. This last procedure has been elegantly presented in several articles by Furihata, Matsuo and collaborators, see e.g. [11, 12, 13, 18, 19, 20, 21], using the concept of discrete variational derivatives. See also the forthcoming monograph [14]. The first part of this paper develops a similar framework that is rather general and allows for an arbitrary number of dependent and independent variables with derivatives of any order. The suggested approach does not require the equations to be discretised in space.

In this paper we consider the class of conservative schemes which are linearly implicit. In fact, this work can be seen as a generalisation of ideas introduced in [19, Section 6]. By linearly implicit we mean schemes which require the solution of precisely one linear system of equations in each time step. This is opposed to fully implicit schemes for which one typically applies an iterative solver that may require a linear system to be solved in every iteration. For standard fully implicit schemes one would typically balance the iteration error in solving the nonlinear system with the local truncation error. However, for conservative schemes the situation is different since exact conservation of the invariant requires that the nonlinear system is solved to machine precision.

It may not, in general, be an easy task to quantify exactly what can be expected of gain in computational cost, if any, when replacing a fully implicit scheme with a linearly implicit one. For illustration we present an example where the KdV equation

$$u_t + u_{xxx} + (u^2)_x = 0 \tag{1.1}$$

is solved on a periodic domain using a fully implicit scheme

$$\frac{U^{n+1} - U^n}{\Delta t} + \frac{U^{n+1}_{xxx} + U^n_{xxx}}{2} + \left( \frac{(U^{n+1})^2 + U^{n+1}U^n + (U^n)^2}{3} \right)_x = 0, \tag{1.2}$$

and a linearly implicit scheme

$$\frac{U^{n+2} - U^n}{2\Delta t} + \frac{U^{n+2}_{xxx} + U^n_{xxx}}{2} + \left( U^{n+1} \frac{U^{n+2} + U^{n+1} + U^n}{3} \right)_x = 0. \tag{1.3}$$

These schemes are derived in Section 6. For space discretisation centered differences are used for both schemes. The schemes are both second order in time, but in our example the linearly implicit multistep scheme has an error constant which is about 3-4 times larger than the fully implicit one-step scheme. In Figure 1.1 we plot the global error versus the number of linear solves for the two schemes. The linearly implicit scheme solves one linear system in each time step. The fully implicit scheme, on the other hand, solves a linear system for each Newton iteration which is repeated to machine precision in each time step. For the largest time step in our experiment this means as much as 561 linear solves per time step. The linear systems in each of the two cases have the same matrix structure, they are both penta-diagonal, and we therefore assume that the cost of solving the linear system is approximately the same for both methods. The $x$-axis in Figure 1.1 can thus be interpreted as a measure of the computational cost in each scheme. The plot shows that for a given global error the linearly implicit scheme is less computationally expensive than the fully implicit scheme.

There are situations in which the results from this example may be less relevant. For instance, the iteration method used in a fully implicit schemes may use approximate versions of the Jacobian for which faster solvers can be applied, therefore the cost of a linear

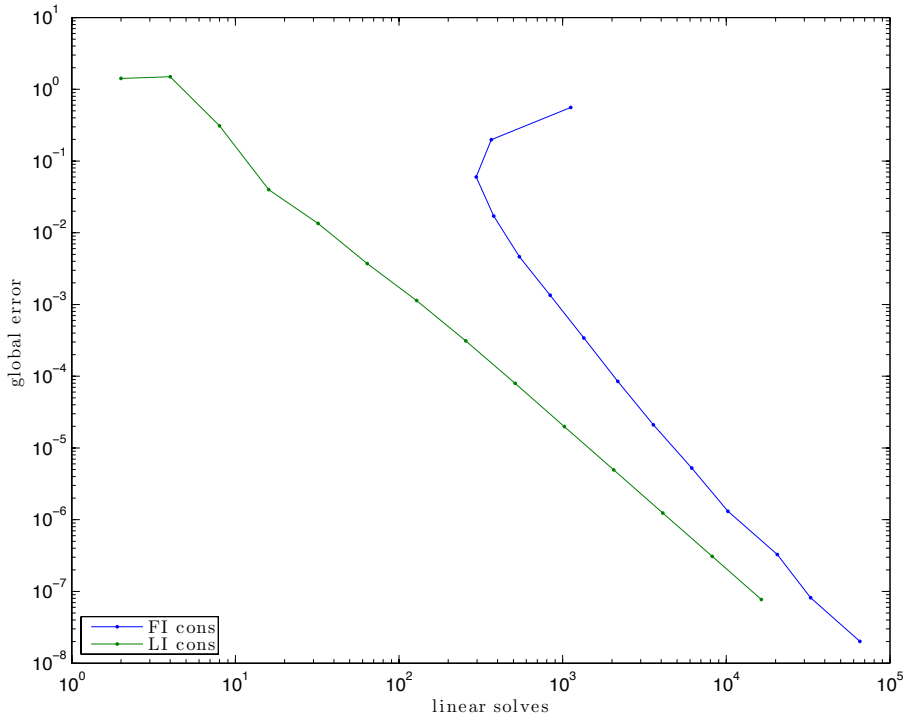



Figure 1.1: The global error versus the number of linear solves for the two schemes (1.2) (FI cons) and (1.3) (LI cons).

solve may not be the same for the two types of schemes. For large time steps, both types of schemes are likely to encounter difficulties, but for slightly different reasons. The fully implicit scheme may experience slow or no convergence at all of the iteration scheme, whereas the linearly implicit scheme may become unstable for time steps over a certain threshold [8]. Thus, we believe that which of the two types of schemes that is preferable depends on the PDE and the circumstances under which it is to be solved. Finally we remind the reader that the main motivation for applying conservative schemes is the long time behavior and stability of the scheme, a facet which is not addressed in the above example since it is common to both types of schemes under consideration. We do however believe that the example provides a convincing argument for the usefulness of linearly implicit versions of conservative integrators. See Section 6 for an example that tests the long time structure preserving properties of these schemes.

In the next section we define the PDE framework that we use. Then, in Section 3 we consider discrete gradient methods and how they can be applied to PDEs. We study in particular the average vector field method by Quispel and McLaren [24] and the discrete variational derivative method by Furihata, Matsuo and coauthors [11, 12, 18, 19, 20, 21]. We develop a framework that works for a rather general class of equations.

The key tools for developing linearly implicit methods for polynomial Hamiltonians are treated in Section 4, introducing the concept of polarisation. There is some freedom in this procedure, and we show through a rather general example term how the choice may significantly affect the stability of the scheme.

We defer the introduction of spatial discretisation until Section 5. This is done mostly in order to keep a simpler notation, but also because our approach concerns conservative time discretisations and is essentially independent of the choice of spatial discretisation. The last section offers some more details on the procedure for constructing schemes and we

give some indication through numerical tests on the long term behaviour of the schemes.

## 2 Notation and preliminaries

We consider integral preserving PDEs written in the form

$$u_t = \mathcal{D}\frac{\delta \mathcal{H}}{\delta u}, \tag{2.1}$$

where

$$\mathcal{H}[u] = \int_\Omega \mathcal{G}[u]\,\mathrm{d}x = \int_\Omega \mathcal{G}((u^\alpha_J))\,\mathrm{d}x \tag{2.2}$$

is the preserved quantity and $\mathcal{D}$ is a skew-symmetric operator that may depend on $u$. We write $\mathrm{d}x = \mathrm{d}x^1 \cdots \mathrm{d}x^d$. We remark in passing that the class of PDEs which can be written in the form (2.1) contains the class of Hamiltonian PDEs, however we do not make the additional assumption that $\mathcal{D}$ satisfies the Jacobi identity [23]. By $(u^\alpha_J)$ we mean $u$ itself, which may be a vector $u = (u^\alpha) \in \mathbb{R}^m$, and all its partial derivatives with respect to all independent variables, $(x^1, \dots, x^d)$, up to and including some order $\nu$. Thus, $J$ is a multi-index, we let $J = (j_1, \dots, j_r)$, where $r = |J|$ the number of components in $J$, and

$$u^\alpha_J = \frac{\partial^r u^\alpha}{\partial x^{j_1} \cdots \partial x^{j_r}}, \quad 0 \le r \le \nu.$$

As in [23], the square brackets in (2.2) are used to indicate that a function depends also on the derivatives of its arguments with respect to the independent variables. In one dimension $d = 1$ and $m = 1$, for example, we have

$$\mathcal{G}[u] = \mathcal{G}((u_J)) = \mathcal{G}(u, \frac{\partial u}{\partial x}, \dots, \frac{\partial^\nu u}{\partial x^\nu}).$$

The variational derivative $\frac{\delta \mathcal{H}}{\delta u}$ is an $m$-vector depending on $u^\alpha_J$ for $|J| \le \nu'$ where $\nu' \ge \nu$. It may be defined through the relation [23, p. 245]

$$\int_\Omega \frac{\delta \mathcal{H}}{\delta u} \cdot \varphi \,\mathrm{d}x = \left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=0} \mathcal{H}[u + \epsilon\varphi], \tag{2.3}$$

for any sufficiently smooth $m$-vector of functions $\varphi(x)$. One may calculate $\frac{\delta \mathcal{H}}{\delta u}$ by applying the Euler operator to $\mathcal{G}[u]$, the $\alpha$-component is given as

$$\left(\frac{\delta \mathcal{H}}{\delta u}\right)^\alpha = \mathbf{E}_\alpha \mathcal{G}[u] \tag{2.4}$$

where

$$\mathbf{E}_\alpha = \sum_{|J| \le \nu} (-1)^{|J|} D_J \frac{\partial}{\partial u^\alpha_J}. \tag{2.5}$$

so that the sum ranges over all $J$ corresponding to derivatives $u^\alpha_J$ featuring in $\mathcal{G}$. We have used total derivative operators,

$$D_J = D_{j_1} \dots D_{j_k}, \quad D_i = \sum_{\alpha, J} \frac{\partial u^\alpha_J}{\partial x^i} \frac{\partial}{\partial u^\alpha_J}.$$

In parts of the paper we refer to Hamiltonians as polynomial, or specifically quadratic. By this we mean that $\mathcal{H}$ is of a form such that $\mathcal{G}$ is a multivariate polynomial in the

indeterminates $u_J^\alpha$, which in the quadratic case is of degree at most two. For example, the KdV equation (1.1) has a polynomial Hamiltonian of order 3

$$\mathcal{H}[u] = \int_\Omega \left( \frac{1}{2} u_x^2 - \frac{1}{3} u^3 \right) \mathrm{d}x.$$

In this case $\mathcal{G} = \mathcal{G}(u, u_x)$ and thus $m = d = \nu = 1$, and we get

$$\frac{\delta \mathcal{H}}{\delta u} = \mathbf{E}\mathcal{G}((u_J)) = \frac{\partial \mathcal{G}}{\partial u} - \frac{\partial}{\partial x}\frac{\partial \mathcal{G}}{\partial u_x} \tag{2.6}$$

$$= -u^2 - u_{xx}. \tag{2.7}$$

We always assume sufficient regularity in the solution and that the boundary conditions on $\Omega$ are such that the boundary terms vanish when doing integration by parts, for example periodic boundary conditions. The operator $\mathcal{D}$ should be skew-symmetric with respect to the $L^2$ inner product

$$\int_\Omega (\mathcal{D}v) w \, \mathrm{d}x = -\int_\Omega v(\mathcal{D}w) \, \mathrm{d}x \quad \forall \, u, w. \tag{2.8}$$

For the KdV case we simply have $\mathcal{D} = \frac{\partial}{\partial x}$.

Furthermore, to be a true Hamiltonian system it should induce a Poisson bracket on the space of functionals as described e.g. in [23, Ch 7.1], meaning that the Jacobi identity must be satisfied. However, the approach presented here only requires $\mathcal{D}$ to be skew-symmetric so that the functional $\mathcal{H}$ is a conserved quantity. In the case that the PDE has more than one Hamiltonian formulation, we may make a choice of which of the integrals to preserve. Our approach does not in general allow for the preservation of more than one Hamiltonian at the same time, for this see the upcoming paper [9].

PDEs such as the wave equation are typically written with $u_{tt}$ appearing on the left hand side, in such cases we double the dimension of $u$ in order to apply the stated framework. For complex equations one may do something similar, splitting either into a real and an imaginary part, or adding in the complex conjugate as a separate variable.

## 3 Discrete gradient and variational derivative methods

Discrete gradient methods for ODEs were introduced by Gonzalez [15]. See also [5], [6], [22], and [16, Chapter V.5]. Recently this idea has been applied to PDEs in the form of the average vector field (AVF) method [4] and in a somewhat more general setting, the discrete variational derivative (DVD) method.

We recall the definition of a discrete gradient as presented for ODEs. If $H : \mathbb{R}^M \to \mathbb{R}$, a discrete gradient is a continuous map $\overline{\nabla} : \mathbb{R}^M \times \mathbb{R}^M \to \mathbb{R}^M$ such that for every $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^M$

$$H(\mathbf{u}) - H(\mathbf{v}) = \overline{\nabla} H(\mathbf{v}, \mathbf{u}) \cdot (\mathbf{u} - \mathbf{v}),$$
$$\overline{\nabla} H(\mathbf{u}, \mathbf{u}) = \nabla H(\mathbf{u}).$$

Since an ODE system preserving $H$ can be written in the form

$$\frac{\mathrm{d}\mathbf{y}}{\mathrm{d}t} = S(\mathbf{y}) \nabla H(\mathbf{y})$$

for some skew-symmetric matrix $S(\mathbf{y})$, one obtains a conservative method simply by defining approximations $\mathbf{y}^n \approx \mathbf{y}(t^n)$ through the formula

$$\frac{\mathbf{y}^{n+1} - \mathbf{y}^n}{\Delta t} = \tilde{S}\,\overline{\nabla} H(\mathbf{y}^n, \mathbf{y}^{n+1}),$$

where $\tilde{S}$, typically allowed to depend on $\mathbf{y}^n$ and $\mathbf{y}^{n+1}$, is some skew-symmetric matrix approximating the original $S$.

There are many possible choices of discrete gradients for a function $H$, see for instance [16, 22]. A particular example is the one used in the AVF method defined as

$$\overline{\nabla}_{\mathrm{AVF}} H(\mathbf{v}, \mathbf{u}) = \int_0^1 \nabla H(\xi\mathbf{u} + (1-\xi)\mathbf{v})\,\mathrm{d}\xi.$$

When applying this approach to PDEs the obvious strategy is to discretise the Hamiltonian $\mathcal{H}[u]$ in space, replacing each derivative by a suitable approximation like e.g. finite differences, to obtain a Hamiltonian $\mathcal{H}_d(\mathbf{u})$ as for ODEs. Similarly, the skew-symmetric operator $\mathcal{D}$ is replaced by a skew-symmetric $M \times M$-matrix $\mathcal{D}_d$ to yield the scheme

$$\frac{\mathbf{u}^{n+1} - \mathbf{u}^n}{\Delta t} = \mathcal{D}_d\,\overline{\nabla}\mathcal{H}_d(\mathbf{u}^n, \mathbf{u}^{n+1}) \tag{3.1}$$

for advancing the numerical solution $\mathbf{u}^n$ at time $t_n$ to $\mathbf{u}^{n+1}$ at time $t_{n+1}$. Examples are worked out for several PDEs in [4].

Furihata, Matsuo and coauthors present a whole framework for discretising PDEs in the variational setting in a series of papers, providing a discrete analogue of the continuous calculus, see for instance [11]. They discretise $\mathcal{G}$ to obtain $\mathcal{G}_d$ using difference operators, and then the integral in $\mathcal{H}$ is approximated by a sum to yield $\mathcal{H}_d$. Then they derive a discrete counterpart to the variational derivative, and finally state the difference scheme in a form which is a perfect analogue to the Hamiltonian PDE system (2.1), letting

$$\frac{\mathbf{u}^{n+1} - \mathbf{u}^n}{\Delta t} = \mathcal{D}_d \frac{\delta \mathcal{H}_d}{\delta(\mathbf{u}^n, \mathbf{u}^{n+1})}.$$

The use of integration by parts in deriving the Euler operator is mimicked by similar summation by part formulas for the discrete case. Their discrete variational derivative is in fact rather similar to a discrete gradient, as it satisfies the relation

$$\mathcal{H}_d(\mathbf{u}) - \mathcal{H}_d(\mathbf{v}) = \langle \frac{\delta \mathcal{H}_d}{\delta(\mathbf{v}, \mathbf{u})}, \mathbf{u} - \mathbf{v} \rangle \tag{3.2}$$

for the discrete $L^2$ inner product.

In the present paper, we focus on the time dimension in most of what follows, thus we shall defer the steps in which $\mathcal{H}$ and thereby $\mathcal{G}$ are discretised in space. But (3.2) makes perfect sense after removing the subscript $d$, replacing $\mathbf{u}$ and $\mathbf{v}$ by functions $u$ and $v$, and the discrete $L^2$ inner product by the continuous one. A discrete variational derivative (DVD) is here defined to be any continuous function $\frac{\delta \mathcal{H}}{\delta(v,u)}$ of $(u^{(\nu)}, v^{(\nu)})$ satisfying

$$\mathcal{H}[u] - \mathcal{H}[v] = \int_\Omega \frac{\delta \mathcal{H}}{\delta(v, u)}(u - v)\,\mathrm{d}x, \tag{3.3}$$

$$\frac{\delta \mathcal{H}}{\delta(u, u)} = \frac{\delta \mathcal{H}}{\delta u}. \tag{3.4}$$

The integrator yields a continuous function $U^n := U^n(x) \approx u(t_n, \cdot)$ for each $t_n$

$$\frac{U^{n+1} - U^n}{\Delta t} = \mathcal{D} \frac{\delta \mathcal{H}}{\delta (U^n, U^{n+1})}. \tag{3.5}$$

By combining (3.3) and (3.5) we see that the method preserves $\mathcal{H}$.

The AVF scheme can of course also be interpreted as a discrete variational derivative method where

$$\frac{\delta \mathcal{H}_{\text{AVF}}}{\delta (v, u)} = \int_0^1 \frac{\delta \mathcal{H}}{\delta u}[\xi u + (1-\xi)v] \, \mathrm{d}\xi. \tag{3.6}$$

The fact that (3.6) verifies the condition (3.3) is seen from the elementary identity

$$\mathcal{H}[u] - \mathcal{H}[v] = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\xi} \mathcal{H}[\xi u + (1-\xi)v] \, \mathrm{d}\xi. \tag{3.7}$$

The derivative under the integral is written

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\xi} \mathcal{H}[\xi u + (1-\xi)v] &= \left. \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \right|_{\varepsilon = 0} \mathcal{H}[v + (\xi + \varepsilon)(u - v)] \\ &= \int_\Omega \frac{\delta \mathcal{H}}{\delta u}[\xi u + (1-\xi)v](u - v) \, \mathrm{d}\mathbf{x} \end{aligned}$$

Now substitute this into (3.7) and interchange the integrals to obtain (3.3).

In most of the cited papers by Furihata, Matsuo and coauthors, the notion of a DVD method is less general than what we just presented, in the sense that the relation (3.2) is not actually used as the defining equation for a discrete variational derivative. Instead the authors present a relatively general format that can be used for discretising $\mathcal{H}$, this format is depending on the class of PDEs under consideration, and they work out the explicit expression for a particular discrete variational derivative. To give an idea of how the format may look like, we briefly review some points from [11] where PDEs of the form (2.2) are considered with $d = m = \nu = 1$ such that $\mathcal{G} = \mathcal{G}(u, u_x)$. $\mathcal{G}$ is assumed to be written as a finite sum

$$\mathcal{G}(u, u_x) = \sum_\ell \alpha_\ell f_\ell(u) g_\ell(u_x). \tag{3.8}$$

where $f_\ell$ and $g_\ell$ are differentiable functions. [1] The form (3.3) is then derived through

$$\begin{aligned} f_\ell(u) g_\ell(u_x) - f_\ell(v) g_\ell(v_x) = &\frac{f_\ell(u) - f_\ell(v)}{u - v} \frac{g_\ell(u_x) + g_\ell(v_x)}{2} (u - v) \\ &+ \frac{g_\ell(u_x) - g_\ell(v_x)}{u_x - v_x} \frac{f_\ell(u) + f_\ell(v)}{2} (u_x - v_x) \end{aligned}$$

followed by an integration by part on the second term. This technique can be extended in any number of ways to allow for more general classes of PDEs. For instance, one may allow for more factors in (3.8), like

$$\mathcal{G}[u] = \sum_\ell \alpha_\ell \prod_J g_{\ell,J}(\partial_J u)$$

and repeated application of the formula $ab - cd = \frac{a+c}{2}(b-d) + \frac{b+d}{2}(a-c)$ to this equation combined with integration by parts will result in a discrete variational derivative.

[1] In [11] the expression is discretised in space and $g_\ell(u_x)$ is replaced by a product $g_\ell^+(\delta_k^+ U_k) g_\ell^-(\delta_k^- U_k)$ where $\delta_k^+$ and $\delta_k^-$ are forward and backward divided differences respectively.

Schemes which are built on this particular type of discrete variational derivative will be called the Furihata methods in the sequel since it was first introduced in [11]. Matsuo et al. extend the method to complex equations in [19], while [12, 18] derive methods for equations with second order time derivatives. Other papers using the discrete variational derivative approach include [13], [21], and [26].

The lack of a general formalism in the papers just mentioned, makes it somewhat difficult to compare the approach to the AVF method and characterise in which cases they lead to the same scheme. Taking for instance the KdV equation (1.1) one easily finds that both approaches lead to the scheme (1.2), however, considering for instance the Hamiltonian

$$\mathcal{H}[u] = \int_\Omega u u_x^2 \,\mathrm{d}x$$

one would obtain two different types of discrete variational derivative in the Furihata method and the AVF method, that is $\frac{\delta \mathcal{H}_F}{\delta(v,u)} \neq \frac{\delta \mathcal{H}_{\mathrm{AVF}}}{\delta(v,u)}$. In some important cases, the Furihata method and the AVF method lead to the same scheme.

**Theorem 3.1.** *Suppose that the Hamiltonian $\mathcal{H}[u]$ is a linear combination of terms of either of the types*

1. $\int_\Omega \partial_J u \cdot \partial_K u \,\mathrm{d}x$ *for multi-indices $J$ and $K$, or*
2. $\int_\Omega g(\partial_J u)\,\mathrm{d}x$ *for differentiable $g : \mathbb{R} \to \mathbb{R}$.*

*Then the AVF and the Furihata methods yield the same scheme*

*Proof.* It suffices to check one general term of each type.

1. We find the variational derivative using (2.3)

$$\frac{\delta \mathcal{H}}{\delta u} = \left((-1)^{|J|} + (-1)^{|K|}\right) \partial_{J+K} u.$$

Inserting the variational derivative into (3.6) gives

$$\frac{\delta \mathcal{H}_{\mathrm{AVF}}}{\delta(v,u)} = \left((-1)^{|J|} + (-1)^{|K|}\right) \partial_{J+K} \left(\frac{u+v}{2}\right).$$

To find the discrete variational derivative of the Furihata method we compute

$$\begin{aligned}\mathcal{H}[u] - \mathcal{H}[v] &= \int_\Omega \partial_J u \cdot \partial_K u - \partial_J v \cdot \partial_K v \,\mathrm{d}x \\ &= \frac{1}{2}\int_\Omega (\partial_J u - \partial_J v)\cdot(\partial_K u + \partial_K v) + (\partial_J u + \partial_J v)\cdot(\partial_K u - \partial_K v)\ \mathrm{d}x\end{aligned}$$

After integration by parts we get

$$\frac{\delta \mathcal{H}_F}{\delta(v,u)} = \left((-1)^{|J|} + (-1)^{|K|}\right) \partial_{J+K} \left(\frac{u+v}{2}\right),$$

and we see that $\dfrac{\delta \mathcal{H}_{\mathrm{AVF}}}{\delta(v,u)} = \dfrac{\delta \mathcal{H}_F}{\delta(v,u)}$.

2. In this case we get

$$\frac{\delta \mathcal{H}}{\delta u} = (-1)^{|J|} \partial_J g'(\partial_J u),$$

so that

$$\begin{aligned}\frac{\delta \mathcal{H}_{\text{AVF}}}{\delta(v,u)} &= (-1)^{|J|} \int_0^1 \partial_J g'(\partial_J(\xi u + (1-\xi)v))\,\mathrm{d}\xi \\ &= (-1)^{|J|} \partial_J \left( \frac{g(\partial_J u) - g(\partial_J v)}{\partial_J u - \partial_J v} \right).\end{aligned}$$

For the Furihata method one would here just compute

$$H[u] - H[v] = \int_\Omega \frac{g(\partial_J u) - g(\partial_J v)}{\partial_J u - \partial_J v} \left(\partial_J u - \partial_J v\right)\,\mathrm{d}x$$

and integration by parts yields $\dfrac{\delta \mathcal{H}_F}{\delta(v,u)} = \dfrac{\delta \mathcal{H}_{\text{AVF}}}{\delta(v,u)}$

□

# 4 Linearly Implicit Difference Schemes

## 4.1 Polarisation

The key to constructing conservative linearly implicit schemes will be to portion out the nonlinearity over consecutive time steps. In effect, this means that we replace the original Hamiltonian $\mathcal{H}$ with an approximate one $H$. We shall call $H$ a polarisation of $\mathcal{H}$ since its definition resembles the way an inner product is derived from a quadratic form. We shall see that the difference scheme resulting from such a polarised Hamiltonian will be a multistep method. The requirements on $H$ are given in the following definition.

**Definition 4.1** (The polarised Hamiltonian)**.** *Given a Hamiltonian $\mathcal{H}[u]$ the polarised Hamiltonian $H$ depends on $k$ arguments, and is:*

- *Consistent*

$$H[u, u, \ldots, u] = \mathcal{H}[u]. \tag{4.1}$$

- *Invariant under any cyclic permutation of the arguments*

$$H[w_1, w_2, \ldots, w_k] = H[w_2, \ldots, w_k, w_1]. \tag{4.2}$$

Polarisations exist for any Hamiltonian, this is asserted by the example

$$H[w_1, w_2, \ldots, w_k] = \frac{1}{k}\left(\mathcal{H}[w_1] + \mathcal{H}[w_2] + \cdots + \mathcal{H}[w_k]\right).$$

We may impose the polarisation directly on the density $\mathcal{G}((u_J^\alpha))$, letting

$$H[w_1, w_2, \ldots, w_k] = \int_\Omega G[w_1, w_2, \ldots, w_k]\,\mathrm{d}x.$$

The conditions (4.1), (4.2) are then inherited as

$$G(u, u, \ldots, u) = \mathcal{G}(u), \qquad G[w_1, w_2, \ldots, w_k] = G[w_2, \ldots, w_k, w_1].$$

In Section 4.2 we will discuss local order of consistency, it will then be convenient to make the stronger assumption that $\mathcal{H}$ and $H$ are at least twice Frechet differentiable. To distinguish from the weaker notion of variational (Gateaux) derivative, we replace $\delta$ by $\partial$, noting that the first derivative in the two definitions are the same when they both exist. We then find from (4.1) and (4.2) that the Frechet derivatives satisfy the relation

$$\frac{\partial \mathcal{H}}{\partial u}[u] = k \, \frac{\partial H}{\partial w_1}[u, \ldots, u]. \tag{4.3}$$

For the second derivatives, we find the identity

$$\frac{\partial^2 H}{\partial w_1 \partial w_j}[u, \ldots, u] = \frac{\partial^2 H}{\partial w_1 \partial w_{k+2-j}}[u, \ldots, u], \quad j = 2, \ldots, \lfloor k/2 \rfloor + 1. \tag{4.4}$$

which is used to compute

$$\frac{\partial^2 \mathcal{H}}{\partial u^2}[u] = \begin{cases} k \left( \dfrac{\partial^2 H}{\partial w_1^2} + 2 \displaystyle\sum_{\ell=2}^{\frac{k+1}{2}} \frac{\partial^2 H}{\partial w_1 \partial w_\ell} \right), & k \text{ odd}, \\ k \left( \dfrac{\partial^2 H}{\partial w_1^2} + 2 \displaystyle\sum_{\ell=2}^{\frac{k}{2}} \frac{\partial^2 H}{\partial w_1 \partial w_\ell} + \frac{\partial^2 H}{\partial w_1 \partial w_{\frac{k}{2}+1}} \right), & k \text{ even}, \end{cases} \tag{4.5}$$

all second derivatives on the right being evaluated at $[u, \ldots, u]$.

### 4.1.1 Polynomial Hamiltonians

The polarisation of polynomial Hamiltonians will be key to constructing linearly implicit schemes. We will now explain in detail how to do this, and we begin with an example term in the integrand $\mathcal{G}[u] = \mathcal{G}(\partial_J^\alpha u)$ depending on just one scalar indeterminate, namely $\mathcal{G}(z) = z^p$ where $z = \partial_J u^\alpha$ for some $(J, \alpha)$ and where $p \leq 4$. This example is important not only as a simple illustration of the procedure, but also because terms of this type are common in many of the Hamiltonians found in physics. As we will see in the next section, it will be natural to use two arguments, $k = 2$, in the polarised Hamiltonian. In fact, we need to restrict ourself to cases with polynomial Hamiltonians for our technique to yield linearly implicit schemes. Then, by using $k \geq \lceil p/2 \rceil$, we can obtain polarised Hamiltonians which are at most quadratic in each argument. We call these quadratic polarisations. We see that if $k = 2$ then cyclic is the same as symmetric $G(u, v) = G(v, u)$, and the possible quadratic polarisations for $p = 2, 3, 4$ are respectively,

$$p = 2: \qquad G(u, v) = \theta \frac{u^2 + v^2}{2} + (1 - \theta) uv \,, \quad \theta \in [0, 1], \tag{4.6}$$

$$p = 3: \qquad G(u, v) = uv \frac{u + v}{2}, \tag{4.7}$$

$$p = 4: \qquad G(u, v) = u^2 v^2. \tag{4.8}$$

Note that for these monomials both the third and fourth order case are uniquely given, but the second order case is not. In Section 4.3 we will consider how the choice of $\theta$ influences the stability of the scheme for a term which appears frequently in PDEs.

We now consider the general case when $\mathcal{G}[u]$ is a multivariate polynomial in $N_\nu$ variables of degree $p$. It suffices in fact to let $\mathcal{G}((u_J^\alpha))$ be a monomial since each term can be treated

separately, for $u \in \mathbb{R}^{N_\nu}$. For a convenient notation, we rename the vector of indeterminates $((u_J^\alpha))$ by using a single index i.e. $u = (u_1, \ldots, u_{N_\nu})$ and write

$$\mathcal{G}(u) = u_{i_1} u_{i_2} \ldots u_{i_p} \tag{4.9}$$

One may use the following procedure for obtaining a quadratic polarisation from (4.9)

1. Group the factors of the right hand side of (4.9) into pairs $z_r = u_{i_{2r-1}} u_{i_{2r}}$ and if $p$ is odd $z_k = u_{i_p}$. Set
$$K(z_1, \ldots, z_k) = z_1 \cdots z_k.$$
Note that there are potentially many ways of ordering the factors in (4.9) which give rise to different polarisations.

2. Symmetrise $K$ with respect to the cyclic subgroup of permutations. Letting the left shift permutation $\sigma$ be defined through $\sigma K(z_1, \ldots, z_k) = K(z_2, \ldots, z_k, z_1)$, we set
$$G(z_1, \ldots, z_k) = \frac{1}{k} \sum_{k=1}^{k} \sigma^{k-1} K(z_1, \ldots, z_k)$$
The resulting $G$ is now both consistent (4.1) and cyclic (4.2).

## 4.2 Linearly Implicit Methods

We may now define the discrete variational derivative for this polarised Hamiltonian as a generalisation of (3.3) and (3.4). We let

$$\frac{\delta H}{\delta(w_1, \ldots, w_{k+1})}$$

be a continuous function of $k+1$ arguments, satisfying

$$H[w_2, \ldots, w_{k+1}] - H[w_1, \ldots, w_k] = \int_\Omega \frac{\delta H}{\delta(w_1, \ldots, w_{k+1})} (w_{k+1} - w_1) \, \mathrm{d}x, \tag{4.10}$$

$$k \, \frac{\delta H}{\delta(u, \ldots, u)} = \frac{\delta \mathcal{H}}{\delta u}. \tag{4.11}$$

Our standard example will be a generalisation of the AVF discrete variational derivative, which we define as

$$\frac{\delta H_{\mathrm{AVF}}}{\delta(w_1, \ldots, w_{k+1})} = \int_0^1 \frac{\delta H}{\delta w_1}[\xi w_{k+1} + (1-\xi) w_1, w_2, \ldots, w_k] \, \mathrm{d}\xi. \tag{4.12}$$

Here the variational derivative on the right hand side, $\frac{\delta H}{\delta w_1}$ is defined as before, considering $H$ as a function of its first argument only, leaving the others fixed. Similar discrete variational derivatives could be derived in a number of different ways. In particular one finds that when the function $H$ is quadratic in all its arguments, the approach used in deriving the Furihata methods would lead to a discrete variational derivative which is identical to that of the AVF-method.

Now we define the polarised discrete variational derivative scheme and prove that, under some assumptions, this scheme is conservative, linearly implicit and has formal order of consistency two.

**Definition 4.2.** *For a Hamiltonian PDE of the form* (2.1)*, let* $H$ *be a polarised Hamiltonian of* $k$ *arguments, satisfying* (4.1) *and* (4.2)*, and suppose that approximations* $U^j$ *to* $u(\cdot, j\Delta t)$ *are given for* $j = 0, \dots, k-1$*.*

- *The polarised DVD (PDVD) scheme is given as*

$$\frac{U^{n+k} - U^n}{k\Delta t} = kD\frac{\delta H}{\delta(U^n, \dots, U^{n+k})}, \quad n \geq 0. \tag{4.13}$$

- $D$ *is a skew-symmetric operator approximating* $\mathcal{D}$*. In* (4.13)*,* $D$ *may depend on* $U^{n+j}$*,* $1 \leq j \leq k-1$ [2]*, and be consistent*

$$D[u, \dots, u] = \mathcal{D}[u]. \tag{4.14}$$

  $D$ *is called cyclic if*

$$D[w_1, w_2, \dots, w_{k-1}] = D[w_2, \dots, w_{k-1}, w_1]. \tag{4.15}$$

- *If the discrete variational derivative is given by* (4.12)*, then the scheme is called the polarised AVF (PAVF) scheme.*

**Theorem 4.3.** *The scheme* (4.13) *is conservative in the sense that*

$$H[U^n, \dots, U^{n+k}] = H[U^0, \dots, U^{k-1}], \quad \forall n \geq 1.$$

*for any polarised Hamiltonian function* $H$*.*

*Proof.* By induction, this is an immediate consequence of (4.10)

□

In a framework as general as the one presented here, it is not possible to present a general analysis for convergence or the order of the truncation error. However, it seems plausible that a necessary condition to obtain a prescribed order of convergence can be derived through a formal Taylor expansion of the local truncation error, we denote this *the formal order of consistency.*

**Theorem 4.4.**

- *The PAVF scheme has formal order of consistency one for any polarised Hamiltonian, and skew-symmetric operator* $D$ *satisfying* (4.14)*.*

- *If in addition* (4.15) *is satisfied, the scheme has formal order of consistency two.*

*Proof.* We show that when the exact solution is substituted into (4.13) where the discrete variational derivative is given by (4.12), then the residual is $\mathcal{O}(\Delta t^2)$. Throughout the proof we assume the existence of Frechet derivatives. Writing, for any $j$, $u^j = u(t_j, \cdot)$ for the exact local solution at $t = t_j$, we get for the left hand side

$$\frac{u^{n+k} - u^n}{k\Delta t} = \partial_t u^n + \frac{k\,\Delta t}{2}\partial_t^2 u^n + \mathcal{O}(\Delta t^2) = \mathcal{D}\frac{\partial \mathcal{H}}{\partial u}\bigg|_{u^n} + \frac{k\Delta t}{2}\left(\frac{\partial \mathcal{D}}{\partial u}(\partial_t u^n)\frac{\partial \mathcal{H}}{\partial u} + \mathcal{D}\frac{\partial^2 \mathcal{H}}{\partial u^2}(\partial_t u^n, \cdot)\right)\bigg|_{u^n} + \mathcal{O}(\Delta t^2). \tag{4.16}$$

[2] $D$ should not depend on $U^{n+k}$ since otherwise the method would no longer be linearly implicit.

Next we expand (4.12) to get

$$\begin{aligned}\frac{\delta H_{\mathrm{AVF}}}{\delta(u^n,\dots,u^{n+k})} &= \int_0^1 \frac{\delta H}{\delta w_1}(\xi u^{n+k}+(1-\xi)u^n, u^{n+1},\dots,u^{n+k-1})\,\mathrm{d}\xi \\ &= \left.\frac{\partial H}{\partial w_1}\right|_{\mathbf{u}} + \left(\left.\frac{k}{2}\,\frac{\partial^2 H}{\partial w_1^2}\right|_{\mathbf{u}} + \Delta t\sum_{j=2}^{k}(j-1)\left.\frac{\partial^2 H}{\partial w_1 \partial w_j}\right|_{\mathbf{u}}\right)(\partial_t u^n,\cdot) + \mathcal{O}(\Delta t^2)\end{aligned}$$

where $\mathbf{u} = (u^n,\dots,u^n)$. Using first (4.4) and then (4.3), (4.5) we find

$$\frac{\delta H_{\mathrm{AVF}}}{\delta(u^n,\dots,u^{n+k})} = \left.\frac{1}{k}\,\frac{\partial \mathcal{H}}{\partial u}\right|_{u^n} + \left.\frac{\Delta t}{2}\,\frac{\partial^2 \mathcal{H}}{\partial u^2}\right|_{u^n}(\partial_t u^n,\cdot) + \mathcal{O}(\Delta t^2). \tag{4.17}$$

Expanding $D$ we get

$$D[u^{n+1},\dots,u^{n+k-1}] = \mathcal{D}[u^n] + \Delta t\sum_{j=1}^{k-1} j\left.\frac{\partial D}{\partial w_j}\right|_{\mathbf{u}}(\partial_t u^n) + \mathcal{O}(\Delta t^2). \tag{4.18}$$

If the cyclicity condition (4.15) holds for $D$, we can simplify (4.18) to obtain

$$\begin{aligned}D[u^{n+1},\dots,u^{n+k-1}] &= \mathcal{D}[u^n] + \left.\frac{k(k-1)\Delta t}{2}\,\frac{\partial D}{\partial w_1}\right|_{\mathbf{u}}(\partial_t u^n) + \mathcal{O}(\Delta t^2) \\ &= \mathcal{D}[u^n] + \left.\frac{k\Delta t}{2}\,\frac{\partial \mathcal{D}}{\partial u}\right|_{u^n}(\partial_t u_n) + \mathcal{O}(\Delta t^2).\end{aligned} \tag{4.19}$$

By substituting into (4.13) the expressions (4.16), (4.17) and (4.19), all terms of zeroth and first order cancel and we are left with $\mathcal{O}(\Delta t^2)$. □

**Theorem 4.5.** *Suppose that the polarised Hamiltonian $H$ is a quadratic polynomial in each of its arguments, then the PAVF scheme is linearly implicit*

*Proof.* Since $H$ is at most quadratic in the first argument, it follows from (2.5) that $\frac{\delta H}{\delta w_1}$ is of degree at most 1 in its first argument, and so we see from (4.12) that

$$\frac{\delta H}{\delta(U^n,\dots,U^{n+k})}$$

is linear in $U^{n+k}$. Since $D$ does not depend on $U^{n+k}$ we conclude that the scheme (4.13) is linearly implicit. □

In some cases one wishes to have time-symmetric numerical schemes, see for example [16]. The numerical scheme (4.13) will in general not be symmetric, however it is not hard to modify the procedure to yield symmetric schemes. One needs to polarise $\mathcal{H}$ such that $H$ is invariant also when the order of its arguments is reversed, it turns out that this can be achieved by symmetrising over the dihedral group rather than just the cyclic one. A similar adjustment must be made for $D$.

We remark that one can construct explicit schemes by using $p$ time steps (as opposed to $k$) in $H$ such that $H$ becomes $p$-linear (as opposed to $k$-quadratic). The rest of the procedure for the explicit case is the same as for the linearly implicit case. Clearly, one expects that explicit schemes will have more severe stability restrictions than the linearly implicit ones.

## 4.3 Stability

In [8] we studied linearly implicit schemes for the cubic Schrödinger equation, and found that two-step schemes can develop a two-periodic instability in time. We also saw that this can be remedied by choosing a different polarisation of the Hamiltonian.

As it turns out, a common case is when the Hamiltonian is a univariate polynomial of order 4 or less. If we polarise this Hamiltonian using two time-steps, we get three linearly independent $H$, corresponding to (4.6), (4.7), and (4.8). The third and fourth order Hamiltonians are uniquely given. However, in the second order case we can choose $\theta \in [0, 1]$ such that the scheme becomes unstable. Since Hamiltonians of the type (4.6) appear in many important PDEs it may be useful to determine which $\theta \in [0, 1]$ lead to unstable schemes.

We choose to study the test equation with Hamiltonian

$$\mathcal{H}[u] = \frac{1}{2}\int_\Omega u_x^2 \,\mathrm{d}x,$$

and a skew-symmetric operator $\mathcal{D}$ which satisfies the eigenvalue equation

$$\mathcal{D}\mathrm{e}^{\mathrm{i}kx} = \mathrm{i}\lambda_k \mathrm{e}^{\mathrm{i}kx}, \quad \lambda_k \in \mathbb{R}$$

for all integers $k$. The Airy equation

$$u_t + u_{xxx} = 0$$

is of this type with $\mathcal{D} = -\partial_x$ and $\lambda_k = -k$. Other equations which have such terms in the Hamiltonian include the nonlinear Schrödinger equation, the linear wave equation, the KdV equation, and the Kadomtsev-Petviashvili equation.

Rewriting (4.6) gives

$$H[v, u] = \frac{1}{2}\int \left(\theta\frac{u_x^2 + v_x^2}{2} + (1-\theta)u_x v_x\right) \mathrm{d}x.$$

And the numerical scheme is

$$\frac{U^{n+2} - U^n}{2\Delta t} = -\mathcal{D}\left(\theta\frac{U_{xx}^{n+2} + U_{xx}^n}{2} + (1-\theta)U_{xx}^{n+1}\right). \tag{4.20}$$

Since this is a linear equation we can use von Neumann stability analysis [7]. We insert the ansatz

$$U^n(x) = \zeta^n \mathrm{e}^{\mathrm{i}kx}$$

to obtain the quadratic equation

$$(1 - \theta\tau\mathrm{i})\zeta^2 - 2(1-\theta)\tau\mathrm{i}\zeta - (1 + \theta\tau\mathrm{i}) = 0, \quad \tau = \lambda_k \Delta t k^2. \tag{4.21}$$

A necessary condition for stability is $|\zeta| \leq 1$ which implies

$$\theta \geq \frac{1}{2} - \frac{1}{2\tau^2}.$$

Assuming that $\{\lambda_k k^2\}_{k\in\mathbb{Z}}$ is unbounded, we must require that $\theta$ is chosen greater than or equal $\frac{1}{2}$. This is exactly the condition found in [8] for the cubic Schrödinger equation. When $\theta \geq \frac{1}{2}$ the roots of (4.21) satisfy $|\zeta_1| = |\zeta_2| = 1$.

In figure 4.1 we solve the Airy equation with the scheme (4.20) using $\theta = 0.5$ and $\theta = 0.49$. We use the initial value $u(x, 0) = \sin(x)$, which, in the exact case, yields the traveling wave solution $u(x, t) = \sin(x + t)$. The $\theta = 0.49$ solution blows up in few time steps, while the $\theta = 0.50$ solution shows no signs of instability. Doing a discrete Fourier transform of the unstable solution we see that the instability starts at high frequencies, that is large $k$, which corresponds to the results shown above.

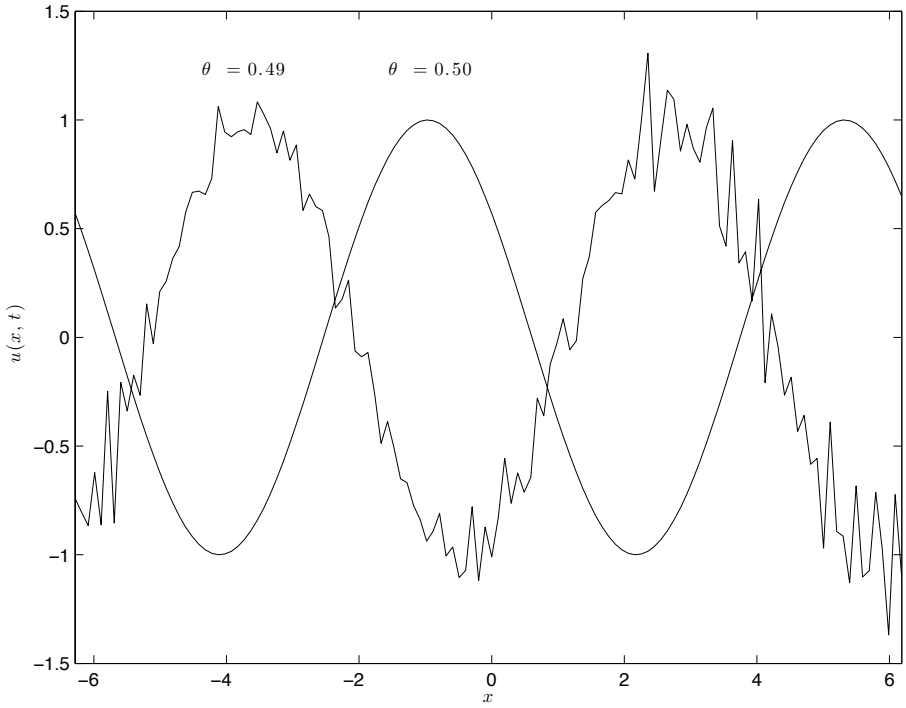



Figure 4.1: The numerical solution of the Airy equation with two different values of $\theta$. The two solutions are shown after $n = 10^6$ time steps ($\theta = 0.5$) and $n = 115$ time steps ($\theta = 0.49$).

# 5 Space discretisation

Until now we have mostly considered the situation where the PDE is discretised in time while remaining continuous in space. The methodology developed in the previous sections apply equally well to systems of ODEs. Arguably, the most straightforward approach is simply to discretise the space derivatives in the Hamiltonian, for instance by finite differences. This leads to

$$\mathcal{H}(u) \longrightarrow \mathcal{H}_d(\mathbf{u}).$$

One also needs to replace the skew-symmetric operator $\mathcal{D}$ by a skew-symmetric matrix $\mathcal{D}_d$. The fully implicit method reviewed in Section 3 is then just the discrete gradient method (3.1), which conserves the discretised Hamiltonian $\mathcal{H}_d(\mathbf{u})$ in every time step.

We consider now finite difference approximations. The function space to which the solution $u$ belongs, is replaced by a finite dimensional space with functions on a grid indexed by $I_g \subset \mathbb{Z}^d$. We use boldface symbols for these functions. Let there be $N_r$ grid points in the space direction $r$ so that $\mathbf{N} = N_1 \cdots N_d$ is the total number of grid points. We denote by $\mathbf{u}^\alpha$ the approximation to $u^\alpha$ on such a grid, and by $\mathbf{u}$ the vector consisting of $(\mathbf{u}^1, \ldots, \mathbf{u}^m)$. We will replace each derivative $u^\alpha_J$ by a finite difference approximation $\delta_J \mathbf{u}^\alpha$, and replace the integral by a quadrature rule. We then let

$$\mathcal{H}_d(\mathbf{u}) = \sum_{\mathbf{i} \in I_g} b_{\mathbf{i}} (\mathcal{G}_d((\delta_J \mathbf{u})))_{\mathbf{i}} \, \Delta x. \tag{5.1}$$

Here $\Delta x$ is the volume (length, area) of a grid cell and $\mathbf{b} = (b_{\mathbf{i}})_{\mathbf{i} \in I_g}$ are the weights in the quadrature rule. The discretised $\mathcal{G}_d$ has the same number of arguments as $\mathcal{G}$, and each input argument as well as the output are vectors in $\mathbb{R}^N$. We have here approximated the

function $u_J^\alpha$ by a difference approximation $\delta_J \mathbf{u}^\alpha$, where $\delta_J : \mathbb{R}^{\mathbf{N}} \to \mathbb{R}^{\mathbf{N}}$ is a linear map. As in the continuous case, we use square brackets, say $F[\mathbf{u}]$, as shorthand for a list of arguments involving difference operators $F[\mathbf{u}] = F(\mathbf{u}, \delta_{J_1}\mathbf{u}, \dots, \delta_{J_q}\mathbf{u})$. We compute

$$\mathcal{H}_d[\mathbf{u}] - \mathcal{H}_d[\mathbf{v}] = \sum_{\mathbf{i}\in I_g} b_{\mathbf{i}} \sum_{J,\alpha} \int_0^1 \left(\frac{\partial \mathcal{G}_d}{\partial \delta_J \mathbf{u}^\alpha}\right)_{\mathbf{i}} [\xi\mathbf{u} + (1-\xi)\mathbf{v}] \mathrm{d}\xi (\delta_J(\mathbf{u}^\alpha - \mathbf{v}^\alpha))\, \Delta x = \langle \frac{\delta \mathcal{H}_d}{\delta(\mathbf{v},\mathbf{u})}, \mathbf{u} - \mathbf{v}\rangle \quad (5.2)$$

where

$$\frac{\delta \mathcal{H}_d}{\delta(\mathbf{v},\mathbf{u})} = \sum_{J,\alpha} \delta_J^T\, B \left( \int_0^1 \frac{\partial \mathcal{G}_d}{\partial \mathbf{u}_J} [\xi \mathbf{u}^\alpha + (1-\xi)\mathbf{v}^\alpha]\, \mathrm{d}\xi \right),$$

$B$ is the diagonal linear map $B = \mathrm{diag}(b_{\mathbf{i}}),\ \mathbf{i} \in I_g$, and the discrete inner product used in (5.2) is

$$\langle \mathbf{u}, \mathbf{v}\rangle = \sum_{\alpha, \mathbf{i}\in I_g} \mathbf{u}_{\mathbf{i}}^\alpha \mathbf{v}_{\mathbf{i}}^\alpha$$

Notice the resemblance between the operator acting on $\mathcal{G}_d$ in (5.2) and the continuous Euler operator in (2.5). Suppose that

1. The spatially continuous method (3.5) (using (3.6)) is discretised in space, using a skew-symmetric $\mathcal{D}_d$ and a selected set of difference quotients $\delta_J$ for each derivative $\partial_J$.

2. Considering (2.4) and (2.5), the choice of discretisation operators $\delta_J$ used in $\partial \mathcal{G}/\partial u_J^\alpha[u]$ is arbitrary, but the corresponding $D_J$ is replaced by the transpose $\delta_J^T$.

In this case, using the same $\mathcal{D}_d$, an identical set of difference operators in discretizing $\mathcal{H}$ (5.1), and choosing all the quadrature weights $b_{\mathbf{i}} = 1$ the resulting scheme is the same.

Letting the $r$th canonical unit vector in $\mathbb{R}^d$ be denoted $\mathbf{e}_r$, we define the most used first order difference operators

$$(\delta_r^+ \mathbf{u})_{\mathbf{i}} = \frac{\mathbf{u}_{\mathbf{i}+\mathbf{e}_r} - \mathbf{u}_{\mathbf{i}}}{\Delta x_r},$$
$$(\delta_r^- \mathbf{u})_{\mathbf{i}} = \frac{\mathbf{u}_{\mathbf{i}} - \mathbf{u}_{\mathbf{i}-\mathbf{e}_r}}{\Delta x_r},$$
$$(\delta_r^{\langle 1\rangle} \mathbf{u})_{\mathbf{i}} = \frac{\mathbf{u}_{\mathbf{i}+\mathbf{e}_r} - \mathbf{u}_{\mathbf{i}-\mathbf{e}_r}}{2\Delta x_r}.$$

These difference operators are all commuting, but only the last one is skew-symmetric. However, for the first two one has the useful identities

$$(\delta_r^+)^T = -\delta_r^-, \qquad (\delta_r^-)^T = -\delta_r^+.$$

Higher order difference operators $\delta_J$ can generally be defined by taking compositions of these operators, in particular we shall consider examples in the next section using the second and third derivative approximations

$$\delta_r^{\langle 2\rangle} = \delta_r^+ \circ \delta_r^-, \quad \delta^{\langle 3\rangle} = \delta^{\langle 1\rangle} \circ \delta^{\langle 2\rangle}.$$

We may now introduce numerical approximations $\mathbf{U}^n$ representing the fully discretized system, the scheme is

$$\frac{\mathbf{U}^{n+1}-\mathbf{U}^n}{\Delta t} = \mathcal{D}_d \frac{\delta \mathcal{H}_d}{\delta(\mathbf{U}^n, \mathbf{U}^{n+1})}.$$

The conservative schemes based on polarisation are adapted in a straightforward manner, introducing a function $H_d[\mathbf{w_1}, \ldots, \mathbf{w_k}]$ which is consistent and cyclic as in (4.1), (4.2), and a skew-symmetric map $D_d$ depending on at most $k-1$ arguments. The scheme is then

$$\frac{\mathbf{U}^{n+k}-\mathbf{U}^n}{k\Delta t} = kD_d \frac{\delta H_d}{\delta(\mathbf{U}^n, \ldots, \mathbf{U}^{n+k})}. \tag{5.3}$$

This scheme conserves the function $H_d$ in the sense that

$$H_d[\mathbf{U}^{n+1}, \ldots, \mathbf{U}^{n+k}] = H_d[\mathbf{U}^0, \ldots, \mathbf{U}^{k-1}], \quad n \geq 0.$$

# 6 Examples

To illustrate the procedures for constructing conservative schemes presented in this paper we consider as an example the generalised Korteweg-de Vries (gKdV) equation

$$u_t + u_{xxx} + (u^{p-1})_x = 0$$

for an integer $p \geq 3$, see for example [25]. The case $p = 3$ is the KdV equation (1.1), the case $p = 4$ is known as the modified KdV equation, and $p = 6$ is sometimes referred to as the mass critical generalised KdV equation. The gKdV can be written as (2.1) with

$$\mathcal{H}[u] = \int_\Omega \left(\frac{1}{2}u_x^2 - \frac{1}{p}u^p\right) \mathrm{d}x, \quad \mathcal{D} = \frac{\partial}{\partial x}.$$

The AVF discrete variational derivative (3.6) gives rise to the fully implicit scheme (3.5)

$$\frac{U^{n+1}-U^n}{\Delta t} + \frac{U_{xxx}^{n+1}+U_{xxx}^n}{2} + \frac{1}{p}\left(\sum_{i=0}^{p-1}(U^{n+1})^{p-1-i}(U^n)^i\right)_x = 0. \tag{6.1}$$

After applying the polarising procedure of Section 4.1 to $\mathcal{H} = \int_\Omega \mathcal{G}\,\mathrm{d}x$ we get $H = \int_\Omega G\,\mathrm{d}x$ which depends on $k = \lceil p/2 \rceil$ arguments

$$G[w_1, \ldots, w_k] = \frac{1}{2k}\sum_{i=1}^k (w_i)_x^2 - \begin{cases} \frac{1}{k}\left(\prod_{j=1}^k w_j^2\right)\left(\sum_{i=1}^k \frac{1}{w_i}\right), & p \text{ odd}, \\ \prod_{j=1}^k w_j^2, & p \text{ even}. \end{cases}$$

After finding the AVF discrete variational derivative from (4.12) we get the linearly implicit PAVF scheme (4.13)

$$\begin{aligned} &\frac{U^{n+k}-U^n}{k\Delta t} + \frac{U_{xxx}^{n+k}+U_{xxx}^n}{2} \\ &\qquad + \begin{cases} \frac{1}{p}\left[\left(\prod_{j=1}^{k-1}(U^{n+j})^2\right)\left(\sum_{i=1}^{k-1}\frac{U^{n+k}+U^n}{U^{n+i}}+1\right)\right]_x = 0, & p \text{ odd}, \\ \frac{1}{2}\left[\left(\prod_{j=1}^{k-1}(U^{n+j})^2\right)\left(U^{n+k}+U^n\right)\right]_x = 0, & p \text{ even}. \end{cases} \end{aligned} \tag{6.2}$$

Notice that $U^{n+k}$ is indeed only appearing as linear terms in this scheme. The schemes (1.2) and (1.3) are found by setting $p = 3$ ($k = 2$) in (6.1) and (6.2), respectively. Following

the procedure of Section 5 one can get a fully discretised scheme by replacing $U$ by $\mathbf{U}$ and the first and third derivative operators by $\delta^{\langle 1\rangle}$ and $\delta^{\langle 3\rangle}$ respectively.

In the Figures 6.1 and 6.2 we compare the conservative methods (1.2) and (1.3) with the fully implicit midpoint method

$$\frac{U^{n+1}-U^n}{\Delta t}+\frac{U^{n+1}_{xxx}+U^n_{xxx}}{2}+\left(\left(\frac{U^{n+1}+U^n}{2}\right)^2\right)_x=0 \tag{6.3}$$

and a naive linearly implicit method

$$\frac{U^{n+1}-U^n}{\Delta t}+\frac{U^{n+1}_{xxx}+U^n_{xxx}}{2}+\left(U^nU^{n+1}\right)_x=0. \tag{6.4}$$

We test the four methods on a soliton solution

$$\Phi(x-ct)=\frac{3c}{2}\mathrm{sech}^2\left(\frac{3\sqrt{c}}{2}(x-ct)\right),\quad c>0$$

using the parameters $c=1$, $x=(-5,5)$, $\Delta x=\frac{10}{32}$ and $\Delta t=0.1$. As an indication of the long time behaviour of the presented schemes we consider to which extent the methods are able to preserve the shape and propagation speed of a soliton solution. We define the two quantities

$$\varepsilon_{\text{shape}}=\min_{\tau}\|U^n-\Phi(\cdot-\tau)\|_2^2 \tag{6.5}$$

and

$$\varepsilon_{\text{distance}}=|\operatorname*{argmin}_{\tau}\|U^n-\Phi(\cdot-\tau)\|_2^2-ct^n|. \tag{6.6}$$

Thus $\varepsilon_{\text{shape}}$ measures the shape error of the numerical solution, and $\varepsilon_{\text{distance}}$ measures the error in the travelled distance of the numerical soliton.

We see in Figure 6.1 the fully implicit schemes preserves the shape better than the linearly implicit ones, and that the conservative schemes perform better than the non-conservative ones. In Figure 6.2 we see that the linearly implicit schemes have a more accurate phase speed than the fully implicit ones. Also note that the linearly implicit schemes are significantly more efficient.

## References


[1] T. B. Benjamin. The stability of solitary waves. *Proc. Roy. Soc. (London) Ser. A*, 328:153–183, 1972.

[2] 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.

[3] T. J. Bridges and S. Reich. Numerical methods for Hamiltonian PDEs. *J. Phys. A*, 39(19):5287–5320, 2006.

[4] E. Celledoni, V. Grimm, R.I. McLachlan, D.I. McLaren, D.R.J. O'Neale, B. Owren, and G.R.W. Quispel. Preserving energy resp. dissipation in numerical PDEs, using the average vector field method. *Preprint*, 2009.

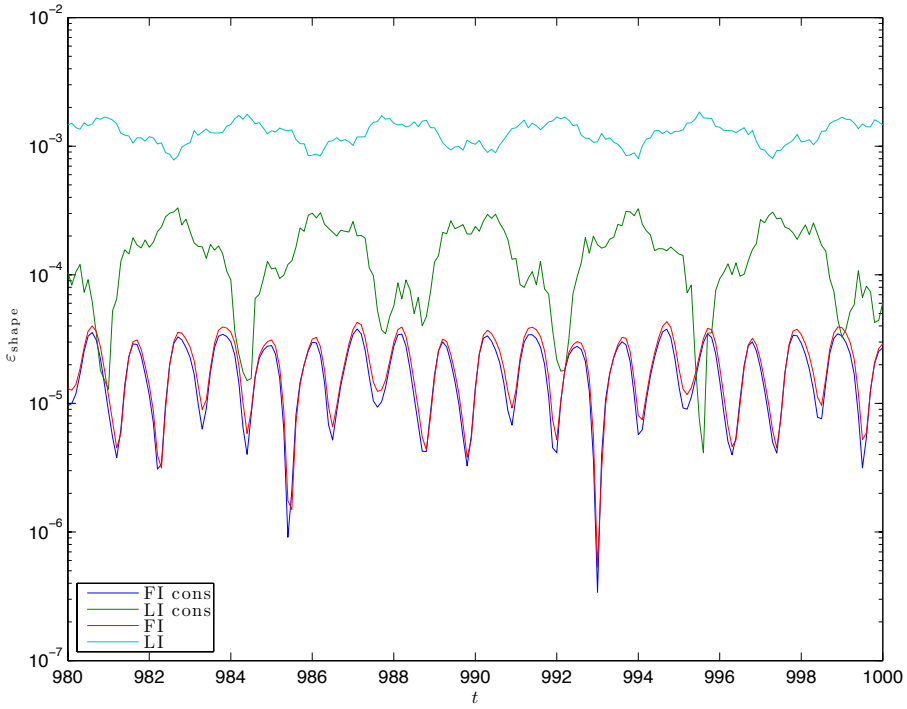



Figure 6.1: The shape error $\varepsilon_{\text{shape}}$ (6.5) for the schemes (1.2) (FI cons), (1.3) (LI cons), (6.3) (FI), and (6.4) (LI). Only the largest times are shown, the plot is similar for smaller $t$.

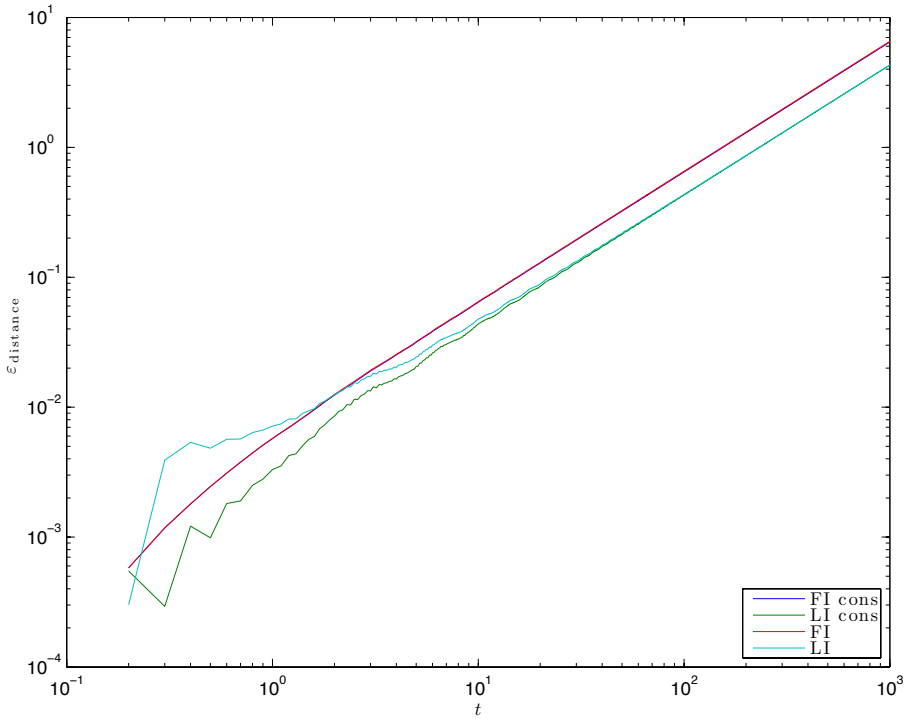



Figure 6.2: The distance error $\varepsilon_{\text{distance}}$ (6.6) for the schemes (1.2) (FI cons), (1.3) (LI cons), (6.3) (FI), and (6.4) (LI). (FI cons) and (FI) are almost indistinguishable in this plot.

[5] E. Celledoni, R.I. McLachlan, D.I. McLaren, B. Owren, G.R.W. Quispel, and W.M. Wright. Energy-preserving Runge-Kutta methods. *Preprint*, 2009.

[6] E. Celledoni, R.I. McLachlan, B. Owren, and G.R.W. Quispel. Energy-preserving integrators and the structure of B-series. *Preprint*, 2009.

[7] J. G. Charney, R. Fjörtoft, and J. von Neumann. Numerical integration of the barotropic vorticity equation. *Tellus*, 2:237–254, 1950.

[8] M. Dahlby and B. Owren. Plane wave stability of some conservative schemes for the cubic Schrödinger equation. *M2AN Math. Model. Numer. Anal.*, 43(4):677–687, 2009.

[9] M. Dahlby, B. Owren, and T. Yaguchi. Preserving multiple first integrals by discrete gradients. *In preparation.*

[10] Z. Fei, V. M. Pérez-García, and L. Vázquez. Numerical simulation of nonlinear Schrödinger systems: a new conservative scheme. *Appl. Math. Comput.*, 71(2-3):165–177, 1995.

[11] D. Furihata. Finite difference schemes for $\partial u/\partial t = (\partial/\partial x)^\alpha \delta G/\delta u$ that inherit energy conservation or dissipation property. *J. Comput. Phys.*, 156(1):181–205, 1999.

[12] D. Furihata. Finite-difference schemes for nonlinear wave equation that inherit energy conservation property. *J. Comput. Appl. Math.*, 134(1-2):37–57, 2001.

[13] D. Furihata. A stable and conservative finite difference scheme for the Cahn-Hilliard equation. *Numer. Math.*, 87(4):675–699, 2001.

[14] D. Furihata and T. Matsuo. *Discrete Variational Method: A Structure-Preserving Numerical Method for Partial Differential Equations.* CRC Press, Boca Raton, in print.

[15] O. Gonzalez. Time integration and discrete Hamiltonian systems. *J. Nonlinear Sci.*, 6(5):449–467, 1996.

[16] E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics.* Springer-Verlag, Berlin, second edition, 2006. Structure-preserving algorithms for ordinary differential equations.

[17] B. Leimkuhler and S. Reich. *Simulating Hamiltonian dynamics*, volume 14 of *Cambridge Monographs on Applied and Computational Mathematics.* Cambridge University Press, Cambridge, 2004.

[18] T. Matsuo. New conservative schemes with discrete variational derivatives for nonlinear wave equations. *J. Comput. Appl. Math.*, 203(1):32–56, 2007.

[19] T. Matsuo and D. Furihata. Dissipative or conservative finite-difference schemes for complex-valued nonlinear partial differential equations. *J. Comput. Phys.*, 171(2):425–447, 2001.

[20] T. Matsuo, M. Sugihara, D. Furihata, and M. Mori. Linearly implicit finite difference schemes derived by the discrete variational method. *Sūrikaisekikenkyūsho Kōkyūroku*, (1145):121–129, 2000. Numerical solution of partial differential equations and related topics (Japanese) (Kyoto, 1999).

[21] T. Matsuo, M. Sugihara, D. Furihata, and M. Mori. Spatially accurate dissipative or conservative finite difference schemes derived by the discrete variational method. *Japan J. Indust. Appl. Math.*, 19(3):311–330, 2002.

[22] R. I. McLachlan, G. R. W. Quispel, and N. Robidoux. Geometric integration using discrete gradients. *R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci.*, 357(1754):1021–1045, 1999.

[23] P. J. Olver. *Applications of Lie groups to differential equations*, volume 107 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1993.

[24] G. R. W. Quispel and D. I. McLaren. A new class of energy-preserving numerical integration methods. *J. Phys. A*, 41(4):045206, 7, 2008.

[25] T. Tao. Why are solitons stable? *Bull. Amer. Math. Soc. (N.S.)*, 46(1):1–33, 2009.

[26] T. Yaguchi, T. Matsuo, and M. Sugihara. Conservative numerical schemes for the Ostrovsky equation. *Journal of Computational and Applied Mathematics*, In Press, Corrected Proof:–, 2009.